

# H0LiCOW – X. Spectroscopic/imaging survey and galaxy-group identification around the strong gravitational lens system WFI 2033–4723

D. Sluse<sup>1,★</sup>, C. E. Rusu<sup>2,3,†</sup>, C. D. Fassnacht<sup>3</sup>, A. Sonnenfeld<sup>4,5</sup>, J. Richard<sup>6</sup>,  
M. W. Auger<sup>7</sup>, L. Coccato<sup>8</sup>, K. C. Wong<sup>4,9</sup>, S. H. Suyu<sup>10,11,12</sup>, T. Treu<sup>13</sup>, A. Agnello<sup>8</sup>,  
S. Birrer<sup>13</sup>, V. Bonvin<sup>14</sup>, T. Collett<sup>15</sup>, F. Courbin<sup>14</sup>, S. Hilbert<sup>16</sup>, L. V. E. Koopmans<sup>17</sup>,  
O. Tihhanova<sup>14</sup>, P. J. Marshall<sup>18</sup>, G. Meylan<sup>14</sup>, A. J. Shajib<sup>13</sup>, J. Annis<sup>19</sup>, S. Avila<sup>20</sup>,  
E. Bertin<sup>21,22</sup>, D. Brooks<sup>23</sup>, E. Buckley-Geer<sup>19</sup>, D. L. Burke<sup>18,24</sup>, A. Carnero  
Rosell<sup>25,26</sup>, M. Carrasco Kind<sup>27,28</sup>, J. Carretero<sup>29</sup>, F. J. Castander<sup>30,31</sup>,  
L. N. da Costa<sup>26,32</sup>, J. De Vicente<sup>25</sup>, S. Desai<sup>33</sup>, P. Doel<sup>23</sup>, A. E. Evrard<sup>34,35</sup>,  
B. Flaugher<sup>36</sup>, J. Frieman<sup>19,37</sup>, J. García-Bellido<sup>20</sup>, D. W. Gerdes<sup>34,35</sup>, D. A. Goldstein<sup>38</sup>,  
R. A. Gruendl<sup>27,28</sup>, J. Gschwend<sup>26,32</sup>, W. G. Hartley<sup>23,39</sup>, D. L. Hollowood<sup>36</sup>,  
K. Honscheid<sup>40,41</sup>, D. J. James<sup>42</sup>, A. G. Kim<sup>43</sup>, E. Krause<sup>44</sup>, K. Kuehn<sup>45</sup>,  
N. Kuropatkin<sup>19</sup>, M. Lima<sup>46,47</sup>, H. Lin<sup>48</sup>, M. A. G. Maia<sup>26,32</sup>, J. L. Marshall<sup>47</sup>,  
P. Melchior<sup>48</sup>, F. Menanteau<sup>27,28</sup>, R. Miquel<sup>49,50</sup>, A. A. Plazas<sup>48</sup>, E. Sanchez<sup>25</sup>,  
S. Serrano<sup>30,31</sup>, I. Sevilla-Noarbe<sup>25</sup>, M. Smith<sup>51</sup>, M. Soares-Santos<sup>52</sup>, F. Sobreira<sup>26,53</sup>,  
E. Suchyta<sup>54</sup>, M. E. C. Swanson<sup>28</sup> and G. Tarle<sup>35</sup>

*Affiliations are listed at the end of the paper*

Accepted 2019 September 1. Received 2019 September 1; in original form 2019 May 17

## ABSTRACT

Galaxies and galaxy groups located along the line of sight towards gravitationally lensed quasars produce high-order perturbations of the gravitational potential at the lens position. When these perturbations are too large, they can induce a systematic error on  $H_0$  of a few per cent if the lens system is used for cosmological inference and the perturbers are not explicitly accounted for in the lens model. In this work, we present a detailed characterization of the environment of the lens system WFI 2033–4723 ( $z_{\text{src}} = 1.662$ ,  $z_{\text{lens}} = 0.6575$ ), one of the core targets of the H0LiCOW project for which we present cosmological inferences in a companion paper. We use the Gemini and ESO-Very Large telescopes to measure the spectroscopic redshifts of the brightest galaxies towards the lens, and use the ESO-MUSE integral field spectrograph to measure the velocity-dispersion of the lens ( $\sigma_{\text{los}} = 250^{+15}_{-21}$  km s<sup>−1</sup>) and of several nearby galaxies. In addition, we measure photometric redshifts and stellar masses of all galaxies down to  $i < 23$  mag, mainly based on Dark Energy Survey imaging (DR1). Our new catalogue, complemented with literature data, more than doubles the number of known galaxy spectroscopic redshifts in the direct vicinity of the lens, expanding to 116 (64) the number of spectroscopic redshifts for galaxies separated by less than 3 arcmin (2 arcmin) from the lens. Using the flexion-shift as a measure of the amplitude of the gravitational perturbation, we identify two galaxy groups and three galaxies that require specific attention in the lens models. The ESO MUSE data enable us to measure the velocity-dispersions of three of these galaxies. These results are essential for the cosmological inference analysis presented in Rusu et al.

\* E-mail: dsluse@uliege.be

† Subaru Fellow.

**Key words:** gravitational lensing: strong – galaxies: groups: general – quasars: individual: WFI 2033–4723.

## 1 INTRODUCTION

The spectroscopic identification of the galaxies located in the environment or along the line of sight (LOS) towards a gravitational lens is one of the important tasks to carry out for deriving an accurate time-delay distance. This is particularly relevant because the lensing cross-section is larger for galaxies residing in rich environment (Fassnacht, Koopmans & Wong 2011; Wong et al. 2018). Depending on their redshift and projected distance from the main lens, galaxies or galaxy groups may significantly perturb the light bending produced by gravitational lensing. The amplitude of the perturbation on the lensed images is larger when the perturber is located in the foreground of the lens, and is maximum at the lens redshift (McCully et al. 2017). The influence on the lensed images also depends on the projected distance of the perturber to the lens. When sufficiently distant in projection to a lens system, galaxy groups (or clusters) produce a uniform convergence at the lens position. This effect can be accounted for in the time-delay distance estimate on a statistical basis, following a methodology similar to the one developed in Rusu et al. (2017). When closer in projection to the lens, galaxies or galaxy groups produce higher order perturbations to the gravitational potential, and therefore must be explicitly included in the lens model; otherwise these perturbations introduce an unknown systematic error. The shift in lensed image positions derived by comparing models with or without the perturber (i.e. the so-called flexion shift) may be used as a criterion to identify objects that need to be included explicitly in the lens model (McCully et al. 2017). For these reasons, it is crucial to obtain spectroscopic and photometric redshifts of the brightest galaxies observable in the field of view (FOV) of a lens system.

The H0LiCOW ( $H_0$  Lenses in COSMOGRAIL’s Wellspring) programme has been initiated with the aim of measuring the Hubble constant  $H_0$  with better than 3.5 per cent accuracy from a small sample of gravitationally lensed quasars with a diversity of observational properties (Suyu et al. 2017). To reach this goal, the programme combines several ingredients: it gathers high-quality data (*Hubble Space Telescope* (HST) imaging, deep images of the FOV, medium resolution spectroscopy of the lens and of nearby galaxies) for each scrutinized system (Suyu et al. 2017), a few-per-cent accuracy measurement of the time delays (Bonvin et al. 2016), and advanced Bayesian lens-modelling techniques (Suyu & Halkola 2010; Suyu et al. 2012; Birrer, Amara & Refregier 2015; Birrer & Amara 2018). An important aspect of the H0LiCOW methodology is that the inferred value of the cosmological parameters (encoded into the so-called time-delay distance) remains blinded until publication. The results are unblinded only when the collaboration considers that all necessary measurements, modelling, and tests have been performed, and then published ‘as is’.

WFI 2033–4723 is part of the H0LiCOW main sample of time-delay lenses. It is a quadruply lensed quasar at redshift  $z_{\text{src}} = 1.662$  lensed by an elliptical galaxy at  $z_{\text{lens}} = 0.6575 \pm 0.001$  (Morgan et al. 2004; Eigenbrod et al. 2006; Sluse et al. 2012; this paper). The minimum and maximum image separation are respectively of  $\Delta\theta_{\text{min}} \sim 0.8$  arcsec and  $\Delta\theta_{\text{max}} \sim 2.5$  arcsec, such that the two brightest images are only barely spatially resolved with ground-based and natural-seeing data, but the two other images are easily photometrically monitored with a 1-m class telescope. Time-delay measurements for the various combinations of image

pairs are presented in Bonvin et al. (2019). When this system was selected to be part of the H0LiCOW sample, the LOS towards the lens was yet to be characterized. An important step forward in the characterization of the lens environment through spectroscopy has been carried out by Momcheva et al. (2015) and Wilson et al. (2016). In particular, Wilson et al. (2016) have confirmed that the lens is part of a massive galaxy group as first suspected by Morgan et al. (2004). Because a proper characterization of the lens environment is crucial to control the systematic errors on  $H_0$ , we have carried out a deeper spectroscopic survey of the FOV of WFI 2033–4723, derived photometric redshifts for the faintest field galaxies, and estimated their stellar masses. Owing to the ESO-MUSE instrument (Bacon et al. 2010), we have been able to carry over a more exhaustive characterization of the galaxies closest to the lens in projection, measuring their redshifts as well as the velocity-dispersions of the lens and of its brightest neighbours. The description and analysis of those new observations, which double the number of spectroscopic measurements for the nearest (in projection) field galaxies, are the main purpose of this paper. They are used to identify and get a proxy on the mass of the main perturbers of the lens potential that need to be explicitly included in the lens-modelling for cosmological inference (Rusu et al. 2019). A joined cosmological inference based on all the lensed systems measured so far by H0LiCOW is presented in Wong et al. (2019).

The paper is structured as follows. We present an overview of the data sets used and of the data reduction process in Section 2. The techniques employed to measure the photometric and spectroscopic redshifts (hereafter photo- $z$  and spec- $z$ , respectively) and stellar masses are presented in Section 3. The methodology used to identify galaxy groups and a list of the groups we identified are described in Section 4. Section 5 quantifies the impact of individual galaxies and galaxy groups on the model. We use the *flexion shift* to flag the systems that require explicit inclusion in the multiplane lens models presented for this system by Rusu et al. (2019). We further measure their velocity-dispersions in Section 6, as this information is included in the lens-modelling presented by Rusu et al. (2019). In addition, we also measure the velocity-dispersion of the lensing galaxy, which is instrumental in reducing the impact of the mass-sheet degeneracy on the lens models. Finally, Section 7 summarizes our main results. In this work, with the exception of the target selection that was based on  $R$ -band magnitude in the Vega system, photometric information comes from multicolour imaging and uses the AB photometric system. For convenience, group radii and masses reported in this work assume a flat  $\Lambda$  cold dark matter cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , and  $\Omega_m = 0.3$ . We stress that this choice has no impact on the group identification as this does not depend on a specific choice of cosmological parameters.

## 2 DATA

Our data set combines multi-object and integral field spectroscopy obtained with Gemini-South and ESO-Paranal observatories, and multiband/deep imaging obtained with the *Spitzer Space Telescope* and the Blanco Telescope, including data from the Dark Energy Survey (DES<sup>1</sup>). The goals of the spectroscopic observations are

<sup>1</sup><https://www.darkenergysurvey.org>

to measure accurate redshifts and identify galaxy groups that need to be explicitly accounted for in the lens model; to measure velocity-dispersions for the massive individual galaxies that are close enough to also require inclusion in the lens model; and to calibrate the photometric redshifts extracted for the galaxies in the imaging data without available spectroscopy. The multiband imaging data complement the spectroscopy, while allowing the measurement of photometric redshifts and stellar masses of galaxies up to a fainter magnitude limit (our set-up yields a typical depth of  $i \sim 23$  mag). Those data are also crucial for the cosmographic analysis as they are instrumental to the estimation of the distribution of convergence at the lens position (see Rusu et al. 2017; Rusu et al. 2019). A summary of the data sets is provided in Table 1.

## 2.1 Imaging

Homogeneous, multiband, large FOV imaging observations are needed in order to achieve a more complete characterization of the environment and the LOS of WFI 2033–4723 than what is possible through targeted spectroscopy. We base our analysis mainly on *grizY*-band DES data included in the Data Release 1 (Abbott et al. 2018) and obtained during 2013 September and 2015 September (2014 September to 2015 October for the  $z$  band). We supplement this with proprietary deep  $u$ -band data observed on 2015 July 21 and 22 (PI. C. E. Rusu) with the Dark Energy Survey Camera (Flaugher et al. 2015) on the Blanco Telescope; VLT/HAWK-I (Pirard et al. 2004; Kissler-Patig et al. 2008) near-infrared data [PI. C. D. Fassnacht, program ID 090.A-0531(A)] observed on 2012 October 12; and with archival IRAC (Fazio et al. 2004) infrared data from the Spitzer Space Telescope (PI. C. S. Kochanek, program ID 20451), observed on 2005 October 20 and 2006 June 4. The characteristics of our data are described in Table 1. We also have WFC3 F160W *HST* imaging data (PI. S. H. Suyu, Program ID 12889) from 2013 April 3 and 4, which is presented in more detail by Suyu et al. (2017) and was only used in this work to check the quality of the star-galaxy classification (see Section 3.1 below).

## 2.2 Spectroscopy

The use of multi-object spectroscopy is optimal to identify group(s) or cluster(s) of galaxies with projected distances of several arcmin from the lens (i.e. typically a few virial radii for groups at  $z > 0.1$ ). In this work, we used the MXU capabilities (multi-object spectroscopy mode with exchangeable laser-cut masks) of the FORS2 instrument (Appenzeller et al. 1998) mounted at the Cassegrain focus of the UT1 (Antu) telescope (PID: 091.A-0642(A), PI: D. Sluse), and the multi-object spectroscopy mode of the Gemini Multi-Object Spectrographs (GMOS; Hook et al. 2004) at the Gemini-South telescope (PID: GS-2013A-Q-2, PI: T. Treu). The instrumental set-up and target selection strategy is similar to the one we used for the lens system HE 0435–1223 and we refer the reader to Sluse et al. (2017) for details on the latter. In brief, we used six masks and the GRIS300V grism + GG435 blocking filter for the FORS2 data, and four masks and the R400 grating with GG455 filter for the GMOS observations. The two instruments allow us to put slits on targets located up to typically 2 arcmin from the lens, and pack approximately 35 long-slits (6 arcsec length) per mask. With a 40 min exposure time per mask (yielding 1 h execution time with overheads), we can measure redshifts of galaxies up to magnitudes  $I \sim 21.5$ . This set-up maximizes the number of observable targets and ensures a large wavelength coverage (typically 4500–8700 Å) to ease redshift detectability. During the observations, the seeing

was always lower than 0.9 arcmin, and airmass  $1 < \sec(z) < 2$ . The FORS2 observations were carried out in service mode between 2013 May 31 and 2014 September 13, while GMOS data were obtained in visitor mode during the nights of 2013 June 03 and 2013 June 06.

The exceptional capabilities of the ESO-MUSE Integral Field Spectrograph, mounted at the Nasmyth B focus of Yepun (ESO-VLT UT4 telescope), offer a natural complement to the multi-object data. Owing to its wide FOV of  $1 \text{ arcmin} \times 1 \text{ arcmin}$ , and a  $0.2 \text{ arcsec} \times 0.2 \text{ arcsec}$  spatial sampling, it allows one to obtain 90 000 simultaneous spectra covering almost the whole optical range (4800–9350 Å) with a resolving power  $R \sim 1800$  to  $\sim 3600$  (i.e.  $2.5 \text{ Å}$  spectral resolution; Richard, 2017). It is therefore perfectly designed to characterize the lens environment on small scales, allowing the measurement of the redshift of the nearest perturbers, and of the velocity-dispersion of the brightest galaxies (including the lensing galaxy). Our observing strategy consist of placing the lensing galaxy close to the centre of the field and obtaining 4 exposures of 600 s, each rotated by 90 degrees with respect to the previous one, and offset by a few spaxels (spatial pixels). The four exposures of 600 s are combined into a single data cube of 2400 s during the data reduction. A first ensemble of 3 combined data cubes has been obtained as part of the Science Verification (SV) programme 60.A-9306(A), on 2014-06-19 and 2014-08-24, allowing us to reach a depth of  $I \sim 25$  mag (continuum emission,  $3\sigma$ ). A second ensemble of six data cubes (Wide Field mode) has been obtained in Service mode on 2016 May 24, 2016 June 29, 2016 July 18, 2016 July 19, 2016 July 20, under programme 097.A-0454(A) (PI: D. Sluse; hereafter P97). Conditions are optimal (i.e. clear sky, seeing better than 0.8 arcsec) only for a fraction of the P97 data. According to the grading scheme established by ESO, two data sets are attributed a grade A (conditions similar to SV data, fulfilled), one a grade B (marginally out of specification), and three a grade C (out of specification). The P97 data are obtained under high moon fraction, and are therefore less deep than the SV data, with depth between  $I \in [21.3, 24.9]$  mag.

## 2.3 Spectroscopy data reduction

We carried out data reduction of the FORS2 and Gemini multi-object spectroscopy data following the same prescriptions as Sluse et al. (2017). The reduction cascade includes the standard steps of spectroscopic data reduction. They are implemented within the ESO *reflex* environment (Freudling et al. 2013) and FORS2 pipeline version 2.2 for FORS data, and through the *gemini-gmos* IRAF<sup>2</sup> subpackage for GMOS data. Of particular relevance for this work is the accuracy at which the wavelength calibration has been performed. For FORS2 data, we used a polynomial of degree  $n = 5$ , which yielded residuals distributed around 0, a RMS of typically  $0.2 \text{ pixels} = 0.66 \text{ Å}$  at all wavelengths and a model accuracy estimated by matching the wavelength solution to the sky lines, to  $0.25 \text{ Å}$ . Comparison of spectra obtained with different instruments confirms the accuracy of the wavelength calibration (see Appendix A).

The MUSE data reduction has been carried out using the MUSE reduction pipeline version 2.0.1 (Weilbacher et al. 2012; Weilbacher et al. 2015). In particular, the standard steps of bias and flat-fielding corrections, wavelength solution, illumination correction, and flux calibration were made for each of the individual exposures with the default parameters of the pipeline. A variance data cube is associated

<sup>2</sup>IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation.

**Table 1.** Overview of the imaging and spectroscopic data set. For spectroscopy, the columns list respectively the instrument used, the number of masks (except for the data obtained with the ESO-MUSE integral field spectrograph), the total number of objects targeted, the approximate resolving power  $R$  of the instrument at central wavelength, the typical wavelength range covered by the spectra (spectra do not always cover the full wavelength range, depending on the exact object location in the field), and the exposure time per mask, or for the full data set in case of ESO-MUSE data. Note that the # of spectra includes duplicated objects. For imaging, the columns list the magnitude depth, filter name, seeing, and exposure time of the data sets used.

Instrument: spectroscopy	# of masks	# of spectra	$R$	$\lambda_1 - \lambda_2$ (Å)	Exp (s)
FORS2	6	236	440	4500–9200	$2 \times 1330$
GMOS	4	130	1000	4400–8200	$4 \times 660$
MUSE <sup>a</sup>	NA	20	1800–3600	4800–9400	$9 \times 4 \times 600$
Imaging <sup>b</sup>	Depth <sup>c</sup> (mag)	Filter	Scale (arcsec)	Seeing (arcsec)	Exp (s)
HAWK-I	$21.5 \pm 0.1$	<i>J</i>	0.1064	0.71	$7 \times 67.5$
HAWK-I	$20.86 \pm 0.08$	<i>H</i>	0.1064	0.71	$3 \times 60$
HAWK-I	$20.76 \pm 0.04$	<i>Ks</i>	0.1064	0.60	$3 \times 60$
DECam	$25.17 \pm 0.06$	<i>u</i>	0.2625	1.16	$65 \times 500$
DES	$24.25 \pm 0.05$	<i>g</i>	0.2625	1.21	$5 \times 90$
DES	$23.8 \pm 0.1$	<i>r</i>	0.2625	0.97	$5 \times 90$
DES	$23.13 \pm 0.08$	<i>i</i>	0.2625	0.81	$6 \times 90$
DES	$22.9 \pm 0.5$	<i>z</i>	0.2625	1.16	$4 \times 90$
DES	$21.4 \pm 0.2$	<i>Y</i>	0.2625	0.92	$7 \times 45$
IRAC	$24.6 \pm 0.3$	3.6	0.600	–	$72 \times 30$
IRAC	$24.0 \pm 0.2$	4.5	0.600	–	$72 \times 30$
IRAC	$22.3 \pm 0.3$	5.7	0.600	–	$72 \times 30$
IRAC	$22.1 \pm 0.3$	7.9	0.600	–	$72 \times 30$
WFC3	$26.4 \pm 0.1$	<i>F160W</i>	0.08	–	26 257

<sup>a</sup>Only 4/9 data sets were obtained within requested observing conditions (graded A by ESO). The others were graded B (1/9) or C (4/9), which means that the seeing was not stable during an observation and/or moon was too close, yielding a high sky level.

<sup>b</sup>The number of exposures for DES data denotes the maximum number of overlaps, as the coverage is not uniform. The pixel scale and exposure time reported for WFC3 characterize the final frame obtained after combining dithered exposures with DrizzlePac (DrizzlePac is a product of the Space Telescope Science Institute, which is operated by AURA for NASA).

<sup>c</sup>We measure  $5\sigma$  detection limits as  $m_{\text{lim}} = \text{ZP} - 2.5 \log(5\sqrt{N_{\text{pix}}}\sigma_{\text{sky}})$ , where ZP is the magnitude zero-point,  $N_{\text{pix}}$  is the number of pixels in a circle with radius 2 arcsec, and  $\sigma_{\text{sky}}$  is the sky-background noise variation. We derive the uncertainty as the standard deviation of the values in 10 empty regions across the frame.

with each data cube produced by the pipeline. It propagates the errors all along the pipeline reduction chain. While the SV data, obtained during dark observing conditions, are little affected by sky subtraction residuals, this is not the case with the P97 mode data. The latter have been post-processed using the Zurich Atmospheric Purge tool (Soto et al. 2016) that improves the sky subtraction by constructing a sky model using principal components analysis. For each data subset, a combined data cube, sampled on a grid of  $0.2 \text{ arcsec} \times 0.2 \text{ arcsec} \times 1.25 \text{ Å}$ , is reconstructed. For the SV data, we combine the three individual data sets, yielding a total exposure time of 7200 s and a median seeing of 1 arcsec. For the P97 data, we tested different combinations of data cubes, minimizing the seeing, amplitude of sky residuals, and optimizing the signal-to-noise ratio (SNR). We find that optimizing the SNR is essential for performing reliable velocity-dispersion measurements of the galaxies. The final datacube for P97 combines 12 exposures, for a total exposure time of 10 800 s.

We note that for FORS2 data, we sometimes included two objects in a slit to maximize the number of observed targets. For that reason, we perform the extraction by fitting a sum of 1D Gaussian profile on each wavelength bin of the rectified 2D spectrum (with  $n = [1, 2]$  depending of the number of objects in the slit). The extraction is

performed on individual exposures of each spectrum, and final 1D spectrum is the result of the coaddition of the wavelength-calibrated extracted spectra of the same target.

### 3 REDSHIFTS AND STELLAR MASSES

#### 3.1 Photometric redshifts and stellar masses

Here we give a brief description of our technique to measure photometric redshifts and stellar masses, which follows the technique described in Rusu et al. (2017). The analysis of the resulting data for estimating the external convergence that is necessary for the cosmological inference will be presented by Rusu et al. (2019).

While the DES and DECam image mosaics cover a very large FOV, the *HST* data cover only the inner  $\sim 2.2 \text{ arcmin} \times 2.6 \text{ arcmin}$  region. In addition, the IRAC and HAWK-I data cover just a few arcminutes around WFI 2033–4723; this is not enough to map the whole area where there is spectroscopic data, but is enough to map the 2 arcmin radius around the lensing system, where structure in the environment and along the LOS has the greatest impact on the lensing model (Collett et al. 2013).

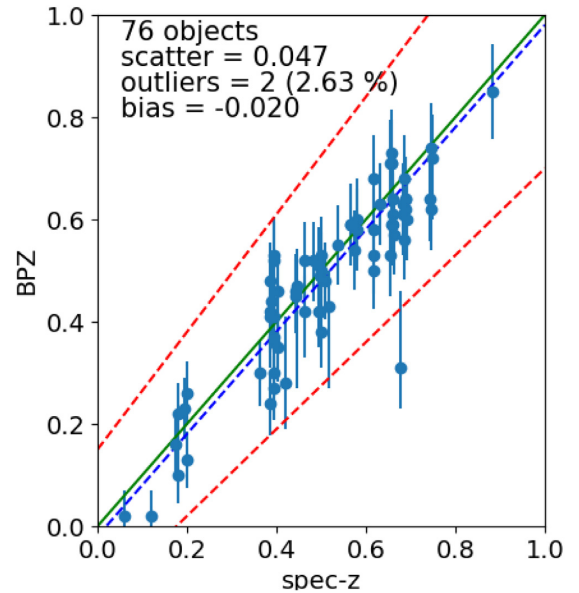


We downloaded cut-outs covering  $4 \text{ arcmin} \times 4 \text{ arcmin}$  around WFI 2033–4723 using the DES cutout service.<sup>3</sup> These consist of *grizY*-band individual exposures that were processed by the DES pipeline (Morganson et al. 2018) to remove the instrumental signature, including bias subtraction, flat-fielding, sky subtraction, artefact masking, and astrometric/photometric calibration. However, at the time when we performed the analysis, master coadded frames were not available. We therefore used *Scamp* (Bertin 2006) to ensure an accurate image registration, and performed image coaddition in each band with *Swarp* (Bertin et al. 2002). We followed similar steps to reduce the DECam *u*-band data (the same instrument used by DES), except that we could not achieve a viable photometric calibration, despite that the observing conditions seemed photometric (but the presence of thin cirrus cannot be excluded).

We reduced the HAWK-I data using the recommended reduction pipeline,<sup>4</sup> in conjunction with *Scamp* and *Swarp*, resampling on to the DES pixel scale, and we calibrated the absolute photometry using bright but unsaturated stars from 2MASS (Skrutskie et al. 2006). In order to enable the measurement of accurate colours between the different filters, we homogenized the shape and size of the point spread function (PSF) by applying suitable convolution kernels. These kernels were computed between two-dimensional Moffat profiles (Moffat 1969) fitted in each band to scaled and stacked bright stars inside the FOV. The resulting PSF full width at half-maximum (FWHM) was  $\sim 1.2 \text{ arcsec}$ .

Our technique to perform object detection and photometric measurements follows that of Erben et al. (2013). For each of the *ugrizYJHKs* bands, *SEXTRACTOR* (Bertin & Arnouts 1996) is run in dual-image mode, where the detection image is the sum of the deepest, best-seeing DES images (*r* and *i*, although we also performed detections in the *i*-band image only), and the measurement images are the PSF-matched images in each of the filters. An additional run performs measurements in the original (i.e. not convolved) *i*-band image. This last run is performed to obtain total magnitudes (*SEXTRACTOR* quantity *MAG\_AUTO*), whereas the previous runs yield accurate colours based on isophotal magnitudes (*MAG\_ISO*). As our resampling and convolutions can produce large noise correlation, which may significantly underestimate the photometric uncertainties measured with *SEXTRACTOR*, we use the technique described in Labbé et al. (2003), Gawiser et al. (2006), and Quadri et al. (2007) to correct for this effect. Finally, we downloaded reduced and photometrically calibrated IRAC data, and we used *T-PHOT* (Merlin et al. 2015) to measure magnitudes matched to the apertures in the DES data, given the much larger pixel scale of the IRAC data and the broader PSF.

We adopt the galaxy–star classification of Hildebrandt et al. (2012). Objects with  $i < 21$  and with size smaller than the PSF are classified as stars. In the range  $21 < i < 23$ , an object is defined as a star if its size is smaller than the PSF and in addition if  $\chi^2_{\text{star}} < 2.0\chi^2_{\text{gal}}$ , where  $\chi^2$  is the best-fitting goodness-of-fit  $\chi^2$  using galaxy and stellar templates. We use both BPZ (Benítez 2000) and EAZY (Brammer, van Dokkum & Coppi 2008) to measure photo-*z*s for the resulting galaxies. Similar to Hildebrandt et al. (2010), we find that the use of currently available mid-IR templates degrade rather than improve the quality of the inferred redshifts. We therefore ignore the IRAC data when estimating redshifts. While the *u*-band data were observed in non-photometric conditions, we solved for its zero-point in a separate run with BPZ by minimizing



**Figure 1.** Comparison of spectroscopic and photometric (BPZ) redshifts for galaxies with robust spectroscopic redshifts within the 120 arcsec radius around the lensing system, based on *ugrizYJHK* photometry. The blue dashed line represents the best-fitting offset, and the green solid line the perfect equality between the two redshift estimates. We define the outliers as data located outside the red dashed line marking  $|z_{\text{spec}} - z_{\text{phot}}|/(1 + z_{\text{spec}}) > 0.15$ . Error bars refer to  $1\sigma$  uncertainties.

the difference between photo-*z*s and spec-*z*s where available. Fig. 1 shows a comparison of the photo-*z*s and spec-*z*s, when the latter exist and are reliable. We also compared the photo-*z*s estimated with BPZ and EAZY. They agree well, with an average scatter of 0.06 and an average outlier fraction [i.e. objects with  $|\Delta z|/(1 + z) > 0.15$ ] of 11 per cent down to the magnitude limit of  $i < 23 \text{ mag}$ .

Finally, since stellar masses are not direct output of BPZ and EAZY, we estimated stellar masses with *LEPHARE* (Arnouts et al. 2002; Ilbert et al. 2010), using galaxy templates based on the stellar population synthesis package of Bruzual & Charlot (2003) with a Chabrier (2003) initial mass function. The stellar mass estimates are performed fixing the redshift to the best fitted photo-*z*. We report the photometry of the  $i < 23 \text{ mag}$  galaxies within 2 arcmin of WFI 2033–4723 in Table C2, and the corresponding redshifts and stellar masses in Table C3. Those tables are also available in electronic form.<sup>5</sup>

In the above, we addressed the galaxies within  $4 \text{ arcmin} \times 4 \text{ arcmin}$  of WFI 2033–4723, where our data provide uniform coverage. For the surrounding FOV of up to 30 arcmin away, we rely on DES data to perform galaxy/star separation and measure photo-*z*s and stellar masses in a similar way. However, instead of performing our own measurements, we rely on total magnitudes provided by the DES pipeline in the form of the *Y3A1\_COADD\_OBJECT\_SUMMARY* table retrieved with *easy-access* (Carrasco Kind et al. 2018). This results in an increased fraction of photo-*z* outliers, from  $\sim 3$  per cent to  $\sim 14$  per cent. We make no effort to improve the extracted colours, as our only use of the resulting quantities is to explore the completeness of our spectroscopic redshifts (see Section 3.3).

<sup>3</sup><https://des.nsa.illinois.edu/easyweb/cutouts>

<sup>4</sup><https://www.eso.org/sci/software/gasgano.html>

<sup>5</sup><http://www.h0licow.org>

**Table 2.** Excerpt of the spectroscopic redshift catalogue. Columns #1–#6 are objects name (=filename of the 1D spectrum), IDs, positions (RA–DEC, J2000), redshifts  $z$ , and their uncertainty  $\sigma_z$ . The last two columns display a quality flag and the object type. The full table is available in electronic form.

Name <sup>a</sup>	ID	RA	DEC	$z$	$\sigma_z$	zQF <sup>b</sup>	Type <sup>c</sup>
FORS_20130531_obj1035	1035	308.404362	−47.392 193	0.5381	0.0002	0	Starburst
FORS_20130531_obj967	967	308.414562	−47.383 083	0.1807	0.0002	0	ETG-Sx
FORS_20130531_obj570	570	308.431962	−47.385 453	0.6174	0.0002	0	Starburst
FORS_20130531_obj445	445	308.470162	−47.401 903	0.4434	0.0002	0	Starburst
FORS_20130531_obj846	846	308.424862	−47.369 983	0.3870	0.0002	0	ETG-Sx

<sup>a</sup>Format: Instrument\_date\_objID, where instrument is *FORS*, *Gemini*, or *MUSE* if the redshift is derived from our survey, and Momcheva if the redshift comes from Momcheva et al. (2015). The ‘date’ in format yyymmdd is the date of observation, or 201508 for objects from Momcheva et al. (2015). This is also the name of the 1D extracted spectrum.

<sup>b</sup>The quality flags zQF = 0/1/2 if the redshift is extracted from this program and 3,4,5,6 refer to objects from Momcheva et al. (2015). zQF = 0 for secure redshift; zQF = 1 for tentative redshift; zQF = 2 for unreliable/unknown redshift; zQF = 3 for data obtained with LDSS-3; zQF = 4 for data obtained with IMACS; zQF = 5 for data obtained with Hectospec; zQF = 6 for NED objects.

<sup>c</sup>Type = ETG-Sx if CaK-H and/or *G* band are detected; Type = Starburst if clear emission lines are observed; Type = M-dwarf for an M-dwarf star; Type = Star for other stellar types; Type = Unknown if no identification could be done or if the spectrum is from an external catalogue.

### 3.2 Spectroscopic redshifts

We followed the methodology of Sluse et al. (2017) for the redshift measurements. Each combined 1D spectrum of an object<sup>6</sup> is cross-correlated with a set of galactic (Elliptical, Sb, only galactic emission lines, quasar) and stellar (*G*, *O*, *M1*, *M8*, *A* spectral types, composite of multiple spectral types) templates using the *xcsao* task, part of the *rvsao* IRAF package (version 2.8.0). Sky regions known to be contaminated by telluric absorption, and/or where sky subtraction is not satisfactory, are masked out. Redshift guesses are derived visually, and refined using the interactive mode of *rvsao*. The redshift from the template providing the highest cross-correlation peak is considered as our final redshift measurement. A flag 0 (secure) / 1 (tentative) / 2 (insecure) is then attached to the spectrum based on the quality of the cross-correlation, signal-to-noise and number of emission/absorption lines detected. The uncertainty on the redshift derived with *xcsao* depends only on the width and peak of the cross-correlation. This error appears to be representative of the statistical uncertainty affecting our measurements, but is smaller than the systematic error as derived by comparing our spectra to literature data (see Appendix A). Unless explicitly stated, the statistical error is used throughout this analysis. It is also the error reported in the final catalogue.

The galaxies detected in the MUSE FOV have been identified automatically using the MUSELET tool, part of the MPDAF package (Piqueras et al. 2017), applied on the combined Science Verification data cubes. Because of the almost dark conditions during the observations, those data allowed us to reach a  $3\sigma$  magnitude limit  $AB = 25.3$  for a point source, i.e. more than 1 mag deeper than any combination of P97 data cubes. The MUSELET tool performs an automatic detection of emission-line features in data cubes by flagging pixels that deviate from the noise (see section 2.2.1 of Drake et al. 2017, for a detailed description). A guess redshift is automatically derived, associating the observed features to brightest multiplets of emission lines detected in galaxy spectra, or to Ly $\alpha$  emission if only one line is detected. The detection of emission in multiple consecutive pixels (along the spectral direction) is used to identify spurious line emission. We visualized the spectra of all

the automatically identified objects to flag obvious artefacts (e.g. sky reduction artefacts that concentrate close to the edge of the FOV). Finally, we compared the catalogue of MUSELET targets to a catalogue of objects detected by running SExtractor on the median data cube (i.e. median along the wavelength direction). This allows the identification of objects that lack emission lines. For all the targets we remeasured the redshift using *rvsao*, following the methodology described above.

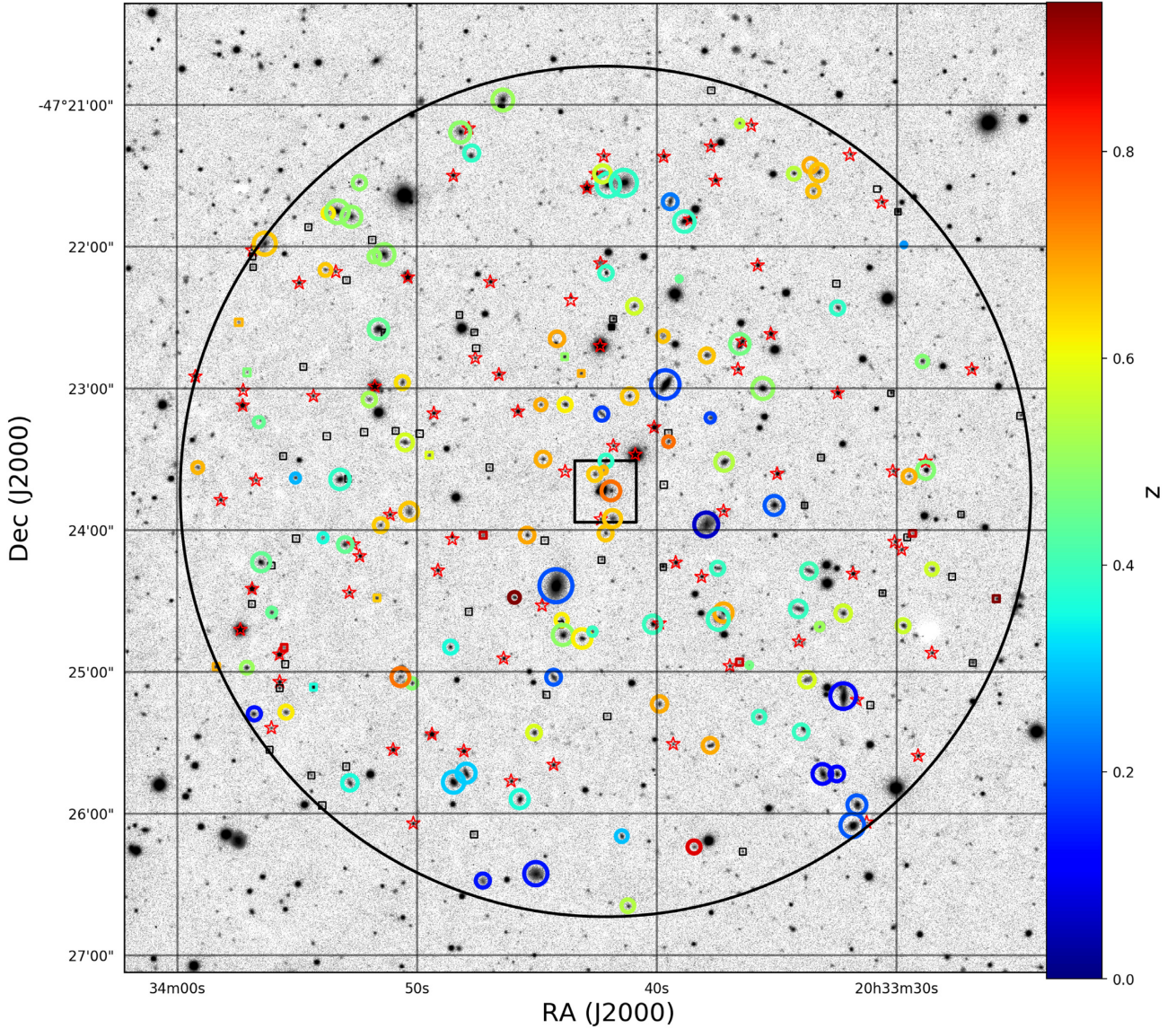
The last step consists in merging the various spectroscopic catalogues into a single one. For each spectrum, an approximate astrometric calibration is deduced based on information recorded in the header of the raw frame. For MXU data, only the position at the centre of the slit is recorded, such that we applied an additional correction based on the object position within the slit and orientation of the laser-cut mask on the sky. Because of the uncertainty of a few arcseconds on the absolute astrometric calibration of the various instruments, and of additional random uncertainties associated with spectral extraction, the astrometric positions between catalogues gathered with different instruments differ by up to 3 arcsec. Since there was a substantial number (i.e.  $>10$ ) of objects in common between pairs of catalogues, we can cross-match catalogues to derive the median astrometric offset (in RA–DEC) ranging from 1.6 to 3.2 arcsec depending of the catalogues considered. Once all the catalogues are virtually matched to the same astrometric system, a new (more robust) cross-correlation can be performed, allowing us to identify duplicates and possible errors in redshift measurements. Objects present in multiple spectroscopic catalogues are found to have compatible redshifts. Instead of combining the multiple measurements, we have decided to keep only the entry with the lowest redshift uncertainty. The final merged catalogue as well as the extracted spectra for the GMOS, FORS2, and MUSE data will be available upon acceptance of the paper in electronic form.<sup>7</sup> The first five lines of the catalogue are displayed in Table 2. Fig. 2 provides an overview of the targets for which spectroscopic information has been gathered within 180 arcsec from WFI 2033–4723. A zoom-in on the galaxies in the central region is displayed in Fig. 3.

The comparison between multiple data sets also provides a good way to flag incorrect redshift measurements, or uncertain ones.

<sup>6</sup>If an object was observed in several masks, redshift measurements were performed independently to avoid introducing biases due to uncertainties in wavelength calibration and/or differences of wavelength coverage.

<sup>7</sup><http://www.holicow.org>





**Figure 2.** Overview of the spectroscopic redshifts obtained from our new and literature data in an FOV of  $\sim 6$  arcmin  $\times$  6 arcmin around WFI 2033–4723 (the black circle delimits a 180 arcsec radius FOV around the lens). Spectroscopically identified stars are marked with a red ‘Star’ symbol, while galaxies are marked with a circle whose size scales with its  $i$ -band magnitude (largest coloured circle correspond to  $i \sim 18.6$  mag, smallest to  $i \sim 23.9$  mag), and colour indicates the redshift (right colour bar). Galaxies that have been targeted but for which no spec- $z$  could be retrieved are shown as open black squares, those with a tentative redshift ( $z_{\text{QF}} = 1$ , see Table 2) with a coloured square (right colour bar). The background frame shows an archival FORS1  $R$ -band combined frame [Prog. ID: 074.A-0563(A)] of 300 s effective exposure time. A zoom on the central region is displayed in Fig. 3.

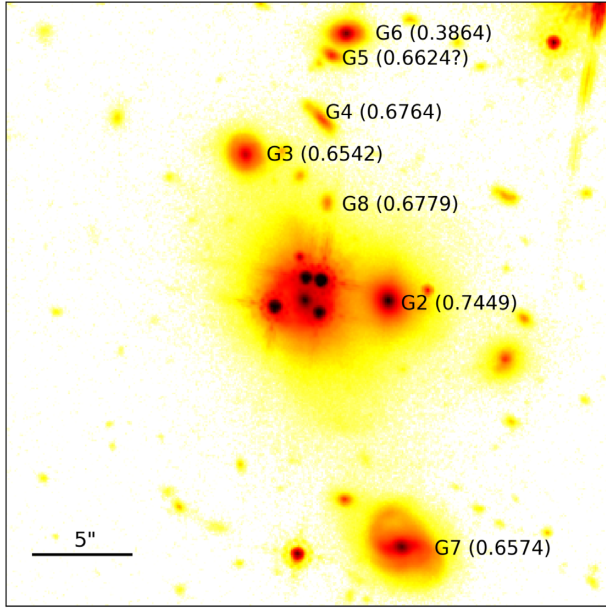
We provide an in-depth cross-comparison of the various data sets used in this work in Appendix A. We found a systematic offset by  $\Delta z = -3.6 \times 10^{-4}$  of the ESO-based data (i.e. FORS and MUSE) compared to GMOS and Momcheva et al. (2015) spectra. While the origin of this offset remains unknown, we have decided to correct the ESO-based measurement for this analysis. In addition, Appendix A lists the four objects for which we suspect a necessary revision of the published redshift.

### 3.3 Completeness of the spectroscopic redshifts

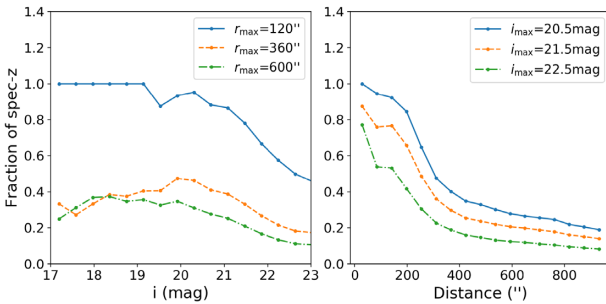
We evaluate the spectroscopic redshift completeness as a function of various criteria by comparing our spectroscopic and photometric catalogues. Fig. 4 displays the completeness of our spectroscopic

catalogue as a function of the limiting magnitude of the sample (fixing the separation to the lens) and of the separation from the lensing galaxy (fixing the limiting magnitude). We see that our completeness is larger than 60 percent at small radius, down to  $i \sim 22.5$  mag. This is similar to the completeness reached for the analysis of HE 0435–1223 (Sluse et al. 2017). However, owing to the MUSE data, we have a higher success in the spectroscopic identification of faint sources located at low projected angular separation from the lens. This is particularly important as those galaxies are most likely to produce high-order perturbations at the lens image position.

Fig. 5 compares the distributions of galaxies (located in projection less than 6 arcmin from the lens) in the spectroscopic and photometric samples, as a function of their median stellar mass

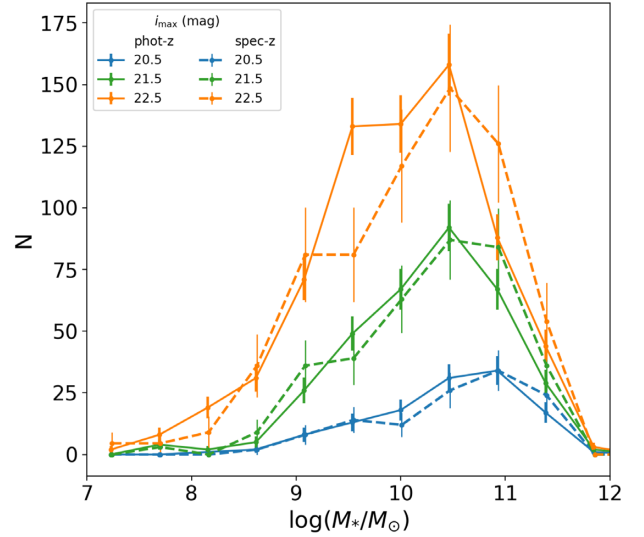


**Figure 3.** Central 30 arcsec  $\times$  30 arcsec region centred on WFI 2033–4723 (matching the central black box in Fig. 2), with galaxy naming scheme G2–G6 following Vuissoz et al. (2008) and G7–G8 are our own designation. North is up, East is left. Redshifts (see Section 3.2) are indicated in parentheses. Insecure redshifts are followed by a question mark.



**Figure 4.** *Left:* Fraction of spectroscopic redshifts of galaxies used in this work (only robust measurements are included) as a function of the maximum  $i$ -band magnitude of the sample, for three different radii  $r_{\max}$  of 2 arcmin (solid-blue), 6 arcmin (dashed-orange), and 10 arcmin (dash-dotted-green). The low apparent completeness for the brightest objects (for  $r_{\max} > 2$  arcmin) is caused by several stars mistakenly classified as galaxies in the photometric catalogue. *Right:* Fraction of spectroscopic redshifts as a function the maximum distance to the lens for three different limiting magnitude [ $i_{\max} = 20.5$  mag (solid-blue);  $i_{\max} = 21.5$  mag (dashed-orange),  $i_{\max} = 22.5$  mag (dot-dashed-green)].

(as derived in Section 3.1). We see that the two distributions agree well, with a slight over-representation of the most massive galaxies ( $M \geq 10^{11} M_{\odot}$ ) in the spectroscopic sample. This is expected as we have a flux limited sample, and more easily measure redshifts of the brightest galaxies. This means that our completeness is the highest for the most massive galaxies, which are also the most likely to perturb the lens gravitational potential. There are no galaxies with  $M \geq 10^{11} M_{\odot}$  within 1 arcmin radius of the lens that are missing spectroscopic redshifts, and only 3 of 12 galaxies if we look up to 2 arcmin separation from the lens. Since those three galaxies are all located at more than 100 arcsec from the



**Figure 5.** Characteristics of the spectroscopic sample for galaxies located less than 6 arcmin from WFI 2033–4723. Number of galaxies as a function of the stellar mass for the photometric (solid) and spectroscopic (dashed) samples for three different cuts in magnitudes  $i_{\max} = (21.5, 22.5, 23.5)$  mag (resp. blue, green, orange). A bin width  $\delta(\log(M/M_{\odot})) = 0.5$  has been considered. To ease legibility, for each magnitude cut, the peak of the distribution of the spectroscopically confirmed galaxies has been normalized by a factor  $n = (2.0, 3.0, 4.5)$  to match the corresponding peak [i.e.  $i_{\max} = (21.5, 22.5, 23.5)$  mag] of the photometric sample. In addition, line-plot instead of bar-plot has been used for clarity.

lens, this ensures that no massive perturber lacks a spectroscopic redshift.

## 4 GALAXY GROUP IDENTIFICATION

The strategy used to identify groups towards WFI 2033–4723 is the same as the one developed by Sluse et al. (2017), building on earlier algorithms implemented in e.g. Wilson et al. (2016). We summarize in Section 4.1 the key aspects of the procedure and refer to Sluse et al. (2017) for more details. Results of the group-finding algorithm are presented in Table 3 and Section 4.2. Discussion and comparison with a previous search for groups towards WFI 2033–4723 are presented in Section 4.3.

### 4.1 Method

After an iterative filtering in redshift space to identify potential group members, an iterative procedure accounting for the 3D position of each galaxy is used to refine group membership and estimate the group velocity-dispersion. In practice, we first select a region of angular radius  $\theta_{\max}$  centred on the lens, bin the redshift catalogue in uniform bins of  $1000 \text{ km s}^{-1}$  (i.e. expected maximal velocity-dispersion of a line-of-sight structure), and identify a redshift peak as a bin of more than  $N$  elements. The operation is repeated after shifting the bins by half the bin width (i.e.  $500 \text{ km s}^{-1}$ ). Then, a first pre-selection (Step #1) of potential group members is performed iteratively for each redshift peak. This is realized by building a core subsample of galaxies that only contains those galaxies separated by less than  $\delta v_{\max}$  from a redshift peak. At each iteration we add galaxies separated by less than  $\delta v_{\max}$  from the average redshift of this core group, and update the group



**Table 3.** Properties of the groups identified in the FOV of WFI 2033–4723. The columns are the group redshift, the number of spectroscopically identified galaxies in the group, the group intrinsic velocity-dispersion (rounded to the nearest  $10 \text{ km s}^{-1}$ ) and  $1\sigma$  standard deviation from bootstrap, the group centroid, bootstrap error on the centroid, projected distance of the centroid to the lens, median flexion shift  $\log(\Delta_3 x (\text{arcsec}))$ , and  $1\sigma$  standard deviation from bootstrapping (Section 5). The last column indicates for which field a peak of more than five galaxies is detected in redshift space. The properties we display correspond to the FOV marked in bold.

ID	$\bar{z}_{\text{group}}$	$N$	$\sigma_{\text{int}}$ (err) ( $\text{km s}^{-1}$ )	RA <sub>ctr</sub> , DEC <sub>ctr</sub> (deg)	err(RA <sub>ctr</sub> , DEC <sub>ctr</sub> ) (arcsec)	$\Delta\theta$ (arcsec)	$\log(\Delta_3 x) \pm \text{err}$ [log(arcsec)]	FOV (arcmin)
b5	0.3060 <sup>a</sup>	8	530 (110)	308.61026100, −47.43226275	145.2, 15.6	469.0	−6.59 ± 0.48	900
a0-1	0.3937 <sup>a</sup>	7	140 (30)	308.40536200, −47.41128275	17.1, 8.8	75.3	−7.25 ± 0.58	360
a0-2	0.3867 <sup>a</sup>	8	100 (20)	308.42486699, −47.36855275	78.1, 17.2	96.9	−7.50 ± 0.57	360
a0-3	0.3999 <sup>a,b</sup>	12	380 (70)	308.34270600, −47.33671275	87.4, 79.4	292.5	−7.01 ± 0.40	360
a2	0.4436	6	150 (40)	308.47654650, −47.39469775	36.3, 21.1	124.2	−7.02 ± 0.54	360
a3	0.4956	13	520 (100)	308.46337200, −47.36336725	61.8, 58.9	147.8	−4.98 ± 0.81	<b>360, 900</b>
a5	0.6588	22	500 (80)	308.43557011, −47.37411275	35.6, 18.6	80.7	−4.70 ± 0.45	<b>360, 900</b>
a8	0.6796 <sup>c</sup>	11	610 (190)	308.42531059, −47.39318538	68.6, 24.3	8.3	−3.75 ± 1.21	360
a9	0.6889 <sup>a</sup>	4	190 (90)	308.41116200, −47.41528275	24.9, 17.5	79.5	−6.42 ± 2.26	360

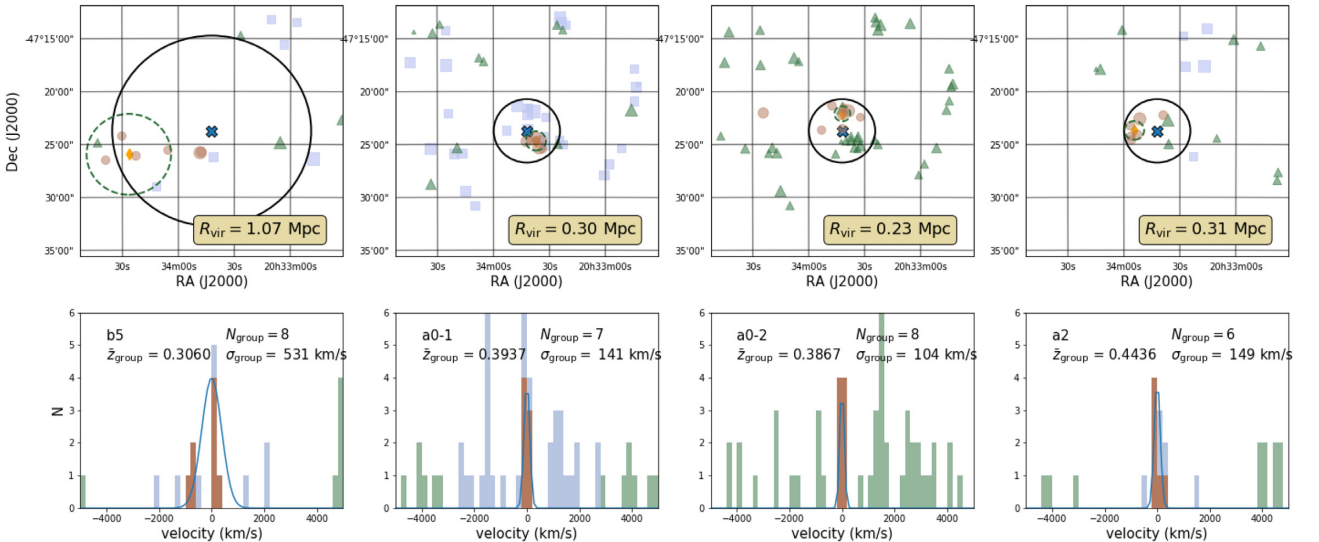
<sup>a</sup>Results from the breakdown of a larger multimodal group candidate.

<sup>b</sup>Likely spurious.

<sup>c</sup>Apparently bimodal but unsuccessful breakdown into sub-group(s).

redshift and velocity-dispersion using a bi-weight estimator. If the new group redshift is found to be more than  $2 \times \delta v_{\text{max}}$  from the estimated redshift, we restrict our search to  $2 \delta v_{\text{max}}$  around the guessed redshift. Our past experience (Sluse et al. 2017) suggests that considering  $\delta v_{\text{max}} = 1500 \text{ km s}^{-1}$  allows one not to miss large distant groups whose effect could be important on the cosmological inference. Based on this filtered galaxy catalogue, we refine group membership (Step #2), accounting for the 3D positions of the galaxies, implementing the method proposed by Wilman et al. (2005). The algorithm selects galaxies located within  $n$  times ( $n = 2$ ) a presumed velocity-dispersion ( $\sigma_{\text{obs}} = 500 \text{ km s}^{-1}$  at the first iteration) along the LOS (i.e. redshift space), and with a maximum aspect ratio between the transverse and radial extension of  $b = 3.5$ .

The maximum extension of the group is deduced from the maximum separation between the group centroid (optionally luminosity weighted) and the candidate galaxy members. The galaxies chosen based on these criteria are used to refine the velocity-dispersion  $\sigma_{\text{obs}}$ . A gapper algorithm (Beers, Flynn & Gebhardt 1990) is used to evaluate  $\sigma_{\text{obs}}$  when there are fewer than 10 galaxies, the dispersion between the velocity measurements when this number drops below 5, and a bi-weight estimator otherwise. This estimate of  $\sigma_{\text{obs}}$  serves as an updated proxy of the velocity-dispersion used to run a new iteration. The algorithm stops when a stable number of group members is found. It also happens that the number of members falls to zero, especially when galaxies are too spread in 3D space (hence not forming a gravitationally linked group). In



**Figure 6.** Main groups identified in the field of WFI 2033–4723: For each redshift (column), the distribution of (rest-frame) velocities of the group galaxies identified spectroscopically is shown (bottom panel) together with a Gaussian of width equal to the intrinsic velocity-dispersion of the group. Bins filled in red correspond to galaxies identified as group members, in blue as interlopers in redshift space, and in green as non-group members. The top panel shows the spatial distribution of the galaxies with a redshift consistent with the group redshift, using the same colour scheme as for the bottom panel. The positions of the lens (group) centroid is indicated with a cross (orange diamond). The size of the symbol is proportional to the brightness of the galaxy, and colour code is the same as for the bottom panel. The solid (dashed) black (green) circles show the field used to identify the peak initial guess for the group redshift (a field of radius  $r \sim 1 \times R_{\text{vir}}$ ). The groups with the largest flexion shifts (and hence, potentially the largest impact on the modelling) are the groups a3, a5 (that includes the lens), and a8 (see Section 5.2, and continued panels of this figure).

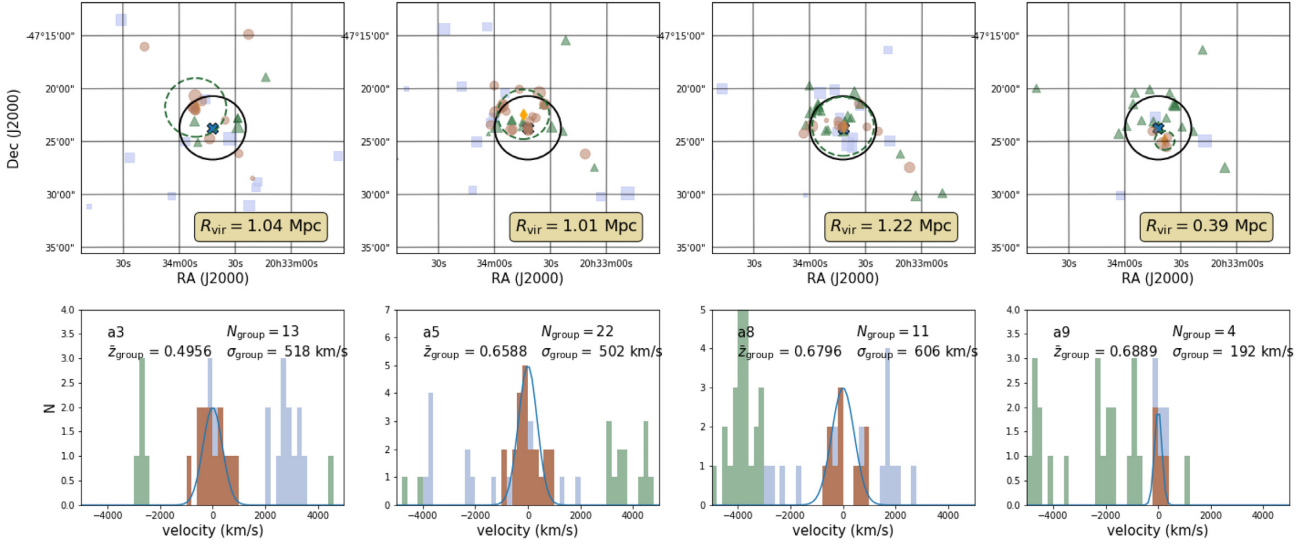


Figure 6. continued.

this situation, no group is associated with the identified redshift peak.

#### 4.2 Results

We have carried out our group search around redshift peaks of  $N \geq 5$  when  $\theta_{\max} = 360$  arcsec, and  $N \geq 10$  when  $\theta_{\max} = 900$  arcsec. The use of two different  $\theta_{\max}$  allows us to avoid missing the identification of a small compact group located close to the lens, if another structure at slightly different redshift (i.e. a few thousands  $\text{km s}^{-1}$ ) is present at larger radius. The difference of cut-off to identify a peak occurs because, at large distance from the lens, we are only interested in identifying the largest groups that could affect directly the lens-modelling. Table 3 lists the properties of the group candidates (Fig. 6). A visual inspection of the automatically detected groups revealed that the algorithm tends to identify multimodal distributions in redshift space as a single large structure, yielding group candidates with characteristics of a galaxy cluster (i.e.  $\sigma_{\text{obs}} \sim 1000 \text{ km s}^{-1}$ ). In such situations, following Muñoz et al. (2013), we run our algorithm around each redshift peak but restricting the search to  $\delta v_{\max} = 500 \text{ km s}^{-1}$  during Step #1, which is also the typical width of the observed modes in the redshift distribution. The drawback of this approach is that the small groups identified this way generally remain unchanged after step #2, even when only very few galaxies fall in projection within one angular virial radius from the group centroid. Consequently, we manually flag those groups as spurious when fewer than two galaxies fall within one angular virial radius from the group centroid. The group centroid is expected to fall close to the brightest galaxy group (Robotham et al. 2011; Shen et al. 2014; Hoshino et al. 2015). Since the use of a luminosity weighting does not improve the match between the group centroid and the brightest galaxy (see Appendix B), we ignore the latter in the remaining parts of our analysis.

#### 4.3 Discussion

Wilson et al. (2016) report the semi-automatic search for groups using a methodology very similar to the one used here. Since our catalogue includes the catalogue used by WIL16, we may expect

to recover their group detection, and/or understand whether some detections were possibly spurious. WIL16 report the automatic detection of five groups towards WFI2033–4723, two of them (at  $\bar{z}_{\text{group}} = 0.1740$  and  $\bar{z}_{\text{group}} = 0.2629$ ) being flagged as uncertain as they are located close to the edge of their FOV. We identified two groups at these redshifts when  $\theta_{\max} = 900$  arcsec, but we removed them from the final list because they contain fewer than 10 members. The three other groups reported by WIL16 are found at a redshift compatible with our groups a3, a5, and a8, but the number of group members is larger by typically 30 per cent in our analysis. The properties of a3 and a5 agree within error bars with our detection, but not a8. In fact, WIL16 report a group of 5 galaxies at  $\bar{z}_{\text{group}} = 0.6838$ , namely located at  $\sim 1200 \text{ km s}^{-1}$  from a8 and  $\sim 1500 \text{ km s}^{-1}$  from a9. Our algorithm also originally identified a group candidate of 20 galaxies with  $\sigma = 1030 \text{ km s}^{-1}$  centred at the same redshift as Wilson’s group (i.e.  $\bar{z}_{\text{group}} = 0.6840$ ), but that group candidate has been broken down into a8 and a9 as the redshift distribution is bimodal, which is not expected in the case of a single group. In addition to automatic detections, WIL16 report three visually identified groups at  $\bar{z}_{\text{group}} = 0.3288$ ,  $\bar{z}_{\text{group}} = 0.3926$ , and  $\bar{z}_{\text{group}} = 0.5100$ , as well as two groups of fewer than five members at  $\bar{z}_{\text{group}} = 0.2151$  and  $\bar{z}_{\text{group}} = 0.3986$ . The groups at  $\bar{z}_{\text{group}} = 0.3926$  and  $\bar{z}_{\text{group}} = 0.3986$  found by WIL16 may be part of the large overdensity of galaxies observed at  $z \sim 0.394$  (i.e. 39 galaxies with  $z \in [0.382, 0.406]$ , or  $\pm 3500 \text{ km s}^{-1}$  from  $z \sim 0.394$ ). The distribution of redshifts in that range is multimodal, suggesting that it is not caused by a massive galaxy cluster.<sup>8</sup> Instead, we identify up to three compact groups (a0–1, a0–2, a0–3), two of them (a0–1 and a0–3) roughly matching the central redshift of the group identified by WIL16. The properties of those groups differ however from those reported by Wilson as our data reveal 18 new galaxies in that redshift range. The other groups reported by WIL16 are found by our algorithm when  $\theta_{\max} = 900$  arcsec, but have been removed because of our choice to only keep groups of at least 10 members for large  $\theta_{\max}$ . The properties of those groups, while not

<sup>8</sup>The visual inspection of  $2 \times 15 \text{ ks}$  archive Chandra ACIS data of WFI2033–4723 shows only point-like sources, but no diffuse emission that would be associated with a galaxy cluster.

identical to those of WIL16 due to the higher completeness of our catalogue, are compatible with the groups of WIL16.

There are two group candidates reported in our work that are absent of WIL16, namely b5 and a2. The group a2 hosts galaxies identified exclusively based on our new data sets. It is therefore expected that WIL16 report no detection at that redshift. On the contrary, six of the eight galaxies identified in b5 were also in WIL16 catalogue. As pointed out in Table 3, the velocity histogram of this group is bimodal, such that its reported properties are likely biased. If split in two, the two sub-groups would miss our threshold of five galaxy members to be classified as a group. Our finding is therefore compatible with the lack of detection by WIL16.

## 5 CONTRIBUTION OF LINE OF SIGHT AND ENVIRONMENT TO THE LENS STRUCTURE

We are interested in identifying the structures (galaxies or galaxy groups) that require explicit modelling in the course of the cosmological inference, but may not be accounted for using a tidal approximation. For that purpose, we need to identify massive galaxies or groups that fall too close in projection to the lens to produce only a uniform perturbation of the main lens gravitational potential over the area covered by the lensed images. As in Sluse et al. (2017), we use the diagnostic proposed by McCully et al. (2014, 2017). The method consists of comparing the shift of the solutions of the lens equation with and without including the flexion produced by the perturber (a single galaxy or a galaxy group). For a point mass, the magnitude of the shift produced by the flexion term, called ‘flexion shift’  $\Delta_{3x}$ , can be written as

$$\Delta_{3x} = f(\beta) \times \frac{(\theta_E \theta_{E,p})^2}{\theta^3}, \quad (1)$$

where  $\theta_E$  and  $\theta_{E,p}$  are the Einstein radii of the main lens and of the perturber, and  $\theta$  is the angular separation on the sky between the lens and the perturber. We define  $f(\beta) = (1 - \beta)^2$  if the perturber is behind the main lens, and  $f(\beta) = 1$  if the galaxy is in the foreground. Here,  $\beta$  is defined for a galaxy at redshift  $z_p > z_d$  as

$$\beta = \frac{D_{dp} D_{os}}{D_{op} D_{ds}}, \quad (2)$$

where the  $D_{ij} = D(z_i, z_j)$  corresponds to the angular diameter distance between redshift  $z_i$  and  $z_j$ , and the subscripts o, d, p, s stand for the observer, deflector, perturber, and source. We explain in the next section how  $\theta_{E,p}$  is determined.

As long as the flexion shift of a galaxy is (much) smaller than the observational precision on the position of the lensed images, its perturbation on the gravitational potential of the main lens can be neglected in the lens model. Based on the simulation results of McCully et al. (2017), we adopt the likely conservative threshold of  $\Delta_{3x} > 10^{-4}$  arcsec, i.e. more than 10 times smaller than the astrometric accuracy of the data used in the cosmological inference analysis. Those authors show that by considering explicitly galaxies or galaxy groups with flexion shift larger than this threshold, we limit the bias on  $H_0$  at the percent level in the cosmological analysis.

### 5.1 Individual galaxies

We first calculate the flexion shift for the individual galaxies in the field of WFI 2033–4723. This requires an estimate of the Einstein radius  $\theta_{E,p}$  of these galaxies. This is achieved in a two-step process.

First, we infer the line-of-sight central velocity-dispersion  $\sigma_{los}$  of each galaxy using the scaling relation from Zahid et al. (2016), and DES-based stellar masses (Section 3.1). This empirical ‘double power-law’ relationship has been derived from a large sample of early-type galaxies at  $z < 0.7$  observed with SDSS, covering the stellar mass range  $\log(M_*/M_\odot) \in [9.5, 11.5]$ . Since no significant modification of the relationship has been found by Zahid et al. (2016) when splitting the sample in different redshift bins, we assume no evolution with redshift. In addition, we assume that this relationship is still valid at the low-mass end of our sample, where  $M_* < 10^{9.5} M_\odot$ . In a few cases, when no accurate multiband photometry was available due to object blending, we fix the stellar mass to  $10^{10.17} M_\odot$ , namely the median stellar mass of the whole sample. We use the relation from Zahid et al. (2016) without regard to the galaxy type. This is a conservative choice as, for the same luminosity, early-types have a larger velocity-dispersion than spirals. Therefore, we may only overestimate the flexion from individual galaxies.

Secondly, we adopt a Singular Isothermal Sphere to convert the velocity-dispersion of the galaxy into its Einstein radius  $\theta_{E,p}$ :

$$\theta_{E,p} = 4\pi \left( \frac{\sigma_{los}}{c} \right)^2 \frac{D_{ps}}{D_{os}}, \quad (3)$$

where  $D_{ps}$  ( $D_{os}$ ) is the angular diameter distance between the perturber  $p$  (resp. the observer  $o$ ) and the source  $s$ . All along the procedure, we use the spectroscopic redshift if available, and the photometric redshift otherwise to calculate distances, together with the stellar mass computed in Section 3.1 at this corresponding redshift. Table 4 lists the 10 galaxies with the largest flexion shifts. Only four of them have a flexion shift  $\Delta_{3x} > 10^{-4}$  arcsec, namely the galaxies labelled G2, G3, G7, and G8 in Fig. 3. Among those galaxies, G8 does not have reliable multiband photometry, and therefore a stellar mass of  $\log(M/M_\odot) = 10.17$  has been assumed. This arbitrary choice may yield a substantial overestimate of the flexion shift. Indeed, this galaxy shows spectroscopic characteristics of a spiral galaxy, and is clearly fainter than G4, another spiral located at about the same redshift than G8, but with a photometric stellar mass of only  $\log(M/M_\odot) \in [8.96, 9.48]$ . Assessing a stellar mass in that range for G8 yields flexion shifts  $\Delta_{3x} \in [3.43 \times 10^{-5}, 2.55 \times 10^{-5}]$  arcsec, well below the threshold above which that galaxy would have a substantial impact on the modelling.

The uncertainty on the flexion shift of each galaxy is derived by quadratically adding the uncertainty originating from the conversion of stellar mass into  $\sigma_{los}$ , with the uncertainty on the stellar mass itself (which is strongly correlated with the photometric redshift, such that we can effectively neglect the redshift uncertainty). More precisely, we calculate the flexion using the 16 and 84 percentile values uncertainty on the velocity-dispersion from fig. 9 of Zahid et al. (2016) to derive the  $1\sigma$  uncertainty originating from the velocity-dispersion; and we calculate the uncertainty originating from the stellar mass by calculating the flexion using the 16 and 84percentile values of the stellar mass. Those two contributions to the error budget yield a typical  $1\sigma$  uncertainty of 0.5 dex on  $\log(\Delta_{3x}/\text{arcsec})$ .

### 5.2 Groups

Because galaxies of a group reside in a common dark matter halo, it is important to assess whether the groups identified in Section 4 need to be explicitly accounted for in the model by attaching a specific mass distribution to their observed centroid. Similarly to the methodology used for the galaxies, we adopt the flexion-shift



**Table 4.** Main characteristics of the 10 galaxies with the largest flexion shift. The first six columns display the galaxy ID (and label used in Fig. 3 if displayed), coordinates (RA, DEC in degrees; ICRS), redshift  $z$ ,  $i$ -band magnitude, and distance to the lensing galaxy (in arcsec). The next three columns provide the logarithm of the flexion shift  $\log(\Delta_3 x / 1 \text{ arcsec})$  for three different percentiles of the posterior distribution, i.e. 16, 50, and 84 per cent (see Section 5.1 for details). Values of the flexion shifts  $\Delta_3 x > 10^{-4}$  arcsec are displayed with bold font to ease the identification of the most prominent perturbers.

ID	RA	DEC	$z$	MAG	dist	$\log(\Delta_3^{16} x)$	$\log(\Delta_3^{50} x)$	$\log(\Delta_3^{84} x)$
501 (G2)	308.424014	-47.395599	0.7449	20.02	3.8	<b>-2.65</b>	<b>-1.99</b>	<b>-1.60</b>
1100 (G8) <sup>a</sup>	308.425195	-47.394358	0.6779	—	4.1	-4.07	<b>-3.51</b>	<b>-3.09</b>
482 (G7)	308.423727	-47.398862	0.6574	20.37	13.0	-4.58	<b>-3.91</b>	<b>-3.54</b>
581 (G3)	308.426804	-47.393648	0.6542	21.19	7.2	-4.98	-4.16	<b>-3.71</b>
1045 (G6)	308.424872	-47.392103	0.3864	21.28	12.3	-5.08	-4.41	-4.01
967	308.414562	-47.383083	0.1807	17.96	52.1	-5.16	-4.60	-4.32
468	308.424933	-47.400595	0.6588	21.01	18.5	-5.47	-4.72	-4.29
574	308.435791	-47.391868	0.6845	20.88	28.1	-5.55	-5.00	-4.63
567	308.420762	-47.384463	0.6574	20.71	41.3	-5.77	-5.28	-4.93
344	308.429001	-47.412962	0.6170	20.05	63.5	-5.68	-5.29	-5.02

<sup>a</sup>Flexion shift likely overestimated due to lack of photometric measurement.

$\Delta_3 x$  (equation 1) as an indicator of the impact of each group on the model. By describing the group as a singular isothermal sphere, we can calculate the group's Einstein radius (equation 3) based on its velocity-dispersion, and hence  $\Delta_3 x$  for each group (Tables 3 and B1).

In order to account for the uncertainty on the group centroid and velocity-dispersion, we have repeated the flexion shift estimate on 1000 bootstrap samples of these quantities. More specifically, we resample with replacement the identified group members (i.e. their position and redshifts) and recalculate the group properties using the resampled members. We calculate the flexion shift for each bootstrap group and estimate the 16 and 84 percentiles based on the bootstrapped distribution. We have conservatively considered that groups for which  $\Delta_3 x > 10^{-4}$  arcsec for more than 5 per cent of the bootstrap samples need to be scrutinized. We discuss below the properties of these groups:

(i) a3 at  $\bar{z}_{\text{group}} = 0.4956$ . The group centroid falls in the vicinity of a subset of five galaxies located within less than 20 arcsec in projection from the lens. One of those galaxies is the second-brightest galaxy of the group candidates, the brightest one being located in the outskirts of the group (in projection).

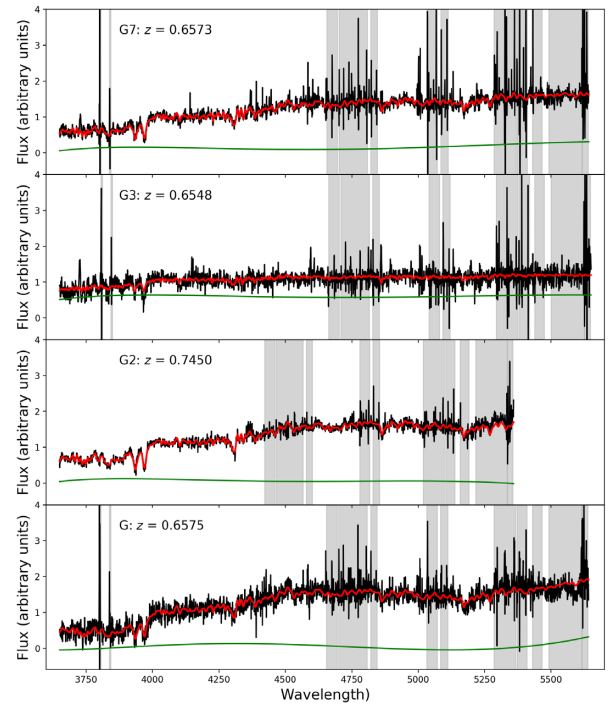
(ii) a5 at  $\bar{z}_{\text{group}} = 0.6588$ : This group hosts the lensing galaxy. The group properties have only a very weak dependence on the weighting scheme used to estimate the group centroid. The latter is distant by about 80 arcsec from several group members, none being the brightest group galaxy.

(iii) a8 at  $\bar{z}_{\text{group}} = 0.6796$ : the distribution in velocity space for this group is very clumpy. This strongly suggests that this group candidate is a spurious detection, as reported in Tables 3 and B1. For that reason, we have decided to discard this group in the lens models used for cosmological inference (Rusu et al. 2019).

In addition to those groups, we have also estimated the flexion shift caused by the group of five galaxies at  $\bar{z}_{\text{group}} = 0.6840$  identified by Wilson et al. (2016). We find  $\Delta_3 x = 1.8 \times 10^{-7}$  arcsec, supporting the small impact of this group candidate on the modelling.

## 6 VELOCITY-DISPERSIONS OF INDIVIDUAL GALAXIES

The velocity-dispersion provides a means of measurement of a galaxy mass. Including this information in the lens-modelling



**Figure 7.** Rest-frame spectra of the galaxies  $G$ – $G2$ – $G3$ – $G7$  (see Fig. 3 for identification) overplotted with the velocity-convolved synthetic stellar population spectrum (red) used to measure the velocity-dispersion. The grey areas display the regions ignored in the velocity-dispersion fitting process due to the presence of known sky absorption or large variability. The green curve is a multiplicative polynomial of the order of 4 used to correct mismatch between the observed spectrum and the synthetic one. The redshift, measured simultaneously with the velocity-dispersion measurement, includes a systematic correction by  $\Delta z \sim -3.6 \times 10^{-4}$  (See Appendix A). Those redshifts agree statistically with those derived in our redshift catalogue (Section 3.2).

allows us to improve the accuracy of the lens models (Treu & Koopmans 2002a,b; Koopmans 2004; Shajib, Treu & Agnello 2018). In addition to the lensing galaxy  $G$ , the three galaxies with the largest flexion shift (i.e.  $G2$ ,  $G3$ ,  $G7$ ; see Fig. 3 and Table 4) are bright enough to enable a velocity-dispersion measurement with MUSE data (Fig. 7).

For that purpose, we use a code that reproduces an observed galaxy spectrum by performing a Bayesian exploration of the stellar population of the galaxy (Auger et al. 2009). More precisely, we model the observed spectrum as a linear combination of stellar spectra multiplied by a sum of orthogonal polynomials (to account for imperfect sky subtraction and uncertainties in the absolute calibration of the spectrum), convolved with a Gaussian kernel to mimic the line-spread-function of the instrument. Contrary to Auger et al. (2009), which uses synthetic stellar spectra, we use an ensemble of real stellar spectra of various types and temperatures (i.e. A0, F2, G0, G5, G8, K1, K2-III stellar types) from the Indo-US spectral library (Valdes et al. 2004). Those spectra cover the rest-frame range [3465–9469] Å, with a constant spectral resolution of 1 Å that correspond to  $\sigma_{\text{template}} \sim 28 \text{ km s}^{-1}$  over the wavelength range considered. The instrumental line-spread-function, derived based on a third order polynomial fit of the spectral resolution with wavelength, is characterized by a median FWHM  $\sim 55 \text{ km s}^{-1}$  over the wavelength range considered. The parameters of the models are the coefficients of the polynomial function that accounts for uncertainties on the flux calibration (nuisance parameters), the coefficients of the linear combination of stellar spectra, the velocity offset compared to the guess redshift, and the velocity-dispersion. The priors are uniform for all these parameters, with a range limited to  $[-350, +350] \text{ km s}^{-1}$  for the velocity, and  $[5, 350] \text{ km s}^{-1}$  for the velocity-dispersion. This methodology, already successfully applied in Auger et al. (2009), Suyu et al. (2010), and Sonnenfeld et al. (2012), is optimized for measuring velocity-dispersion of spectra with  $\text{SNR} \geq 10$  per pixel.

A brute force extraction of the spectra of *G* and *G2* yields significant contamination of the galaxy spectra by the lensed quasar images, precluding a robust velocity-dispersion measurement. It is therefore necessary to remove the quasar flux and local sky residual prior to the velocity-dispersion measurement. For that purpose, we first model each slice of the data cube containing the lens system and *G2* using a 2D model of the light distribution. This model is constituted of the sum of several components: (1) a de Vaucouleurs light profile, convolved with a PSF model; (2) 4-point-like sources whose relative positions are held fixed to the *HST* positions; and (3) a spatially uniform sky background. The code, already successfully used to model ESO-SINFONI IFS data of gravitationally lensed quasars (Braibant et al. 2014; Sluse et al. 2015), uses the MPFIT implementation of the Levenberg–Marquardt least-square optimization algorithm (Markwardt 2009). The lowest residuals are obtained using a symmetric Moffat profile for the PSF. The extraction of the galaxy spectrum is performed within a fixed aperture on the data cube after subtraction of the lensed images and sky models. The associated error is calculated by summing the variance of each spaxel in the aperture.

The velocity-dispersion measurement steps take as input the spectrum of the galaxy and its associated variance. It is necessary to mask regions of the spectra affected by telluric sky absorption and/or residual sky background not perfectly removed by the reduction procedure, as those features may be mistakenly attributed to stellar features (in a complex way that depends on the object’s redshift). For that purpose, we have performed the measurements using different masking schemes (see Appendix C for details). Multiple aperture radii have been tested for the extraction, and we choose 4 pixels radius (i.e. a square of 9 pixels = 1.8 arcsec side-length) as the best compromise between an aperture too small compared to the seeing, and an aperture too large such that the uncertainty on the estimate of the sky subtraction and PSF modelling of nearby targets

**Table 5.** Median velocity-dispersion and 68.4 per cent CI of four galaxies in the FOV of WFI 2033–4723 (see Fig. 2). Measurements are performed within a square aperture of  $3 \times 3$  pixels (1.8 arcsec) centred on the galaxy. The last column indicates which MUSE data set has been used for the velocity-dispersion measurement.

Name	RA, DEC (deg)	$\sigma$ ( $\text{km s}^{-1}$ )	CI on $\sigma$ ( $\text{km s}^{-1}$ )	Note
G	(308.425 58, −47.395 47)	250	229, 267	SV
G2	(308.424 02, −47.395 60)	232	222, 243	P97
		218	213, 222	SV
		223	215, 237	P97 + SV
G3	(308.426 80, −47.393 65)	79	60, 102	SV
G7	(308.423 73, −47.398 86)	166	160, 173	SV

(quasar images and/or nearby galaxies) contribute to a large fraction of the integrated galaxy flux and introduce a large systematic error on its flux. The final velocity-dispersion measurement (Table 5, Fig. 7) results from the marginalization of the probability distribution function obtained for the different masking schemes, and three choices of polynomial order (i.e. order 3, 4, 5). The confidence interval is defined as the region centred on the median and including 68.4 per cent of the probability distribution.

The seeing has been estimated by fitting a Gaussian profile independently for each wavelength slice, on three field stars. In that process, we have ignored spectral slices masked out for measuring the velocity-dispersion. In addition, the seeing has also been estimated on the quasar images when modelling the lens-system luminosity profile. We observed an apparent bias in the FWHM measurement caused by sky residuals (FWHM agree between the different stars better for SV data than for P97 data, and the agreement is better at redder wavelengths). Therefore, we use the FWHM of the brightest star as our proxy of the PSF width. The latter agrees with the FWHM derived from the quasar but is systematically larger by 10 per cent. Because of the more complex measurement of the quasar FWHM, we decided to choose the star-based FWHM as our proxy of the seeing. We measured a median and scatter (along the wavelength direction) of the seeing:  $\text{FWHM} = 1.06 \pm 0.1$  arcsec for the SV data and  $\text{FWHM} = 0.98 \pm 0.25$  arcsec for the P97 data.

## 7 SUMMARY

In the framework of inferring  $H_0$  from the time-delay lens system WFI 2033–4723 (Bonvin et al. 2019; Rusu et al. 2019; Wong et al. 2019), we have performed a detailed characterization of the environmental properties of this system, with the following immediate objectives: (1) identify individual galaxies and galaxy groups susceptible to produce high-order perturbation to the lens potential and therefore requiring to being explicitly included in the lens models and estimate their masses; (2) derive redshift proxies (i.e. spec-zs or photo-zs) for all the galaxies in the field of view to enable a statistical estimate of the convergence associated with the galaxies along the LOS; (3) measure the velocity-dispersion of the main lensing galaxy for self-consistency of the mass modelling.

To reach these goals we have measured photo-zs, spec-zs, and inferred stellar masses for most of the galaxies up to  $\approx 4$  arcmin from the lens, down to  $i \approx 23.0$  mag. We have used deep multicolour imaging as well as multi-objects and integral-field spectroscopy. In particular, we used *grizY* imaging from the Dark Energy Survey, proprietary *u* band obtained with DECAM,

near-infrared HAWK-I (*JHKs*) and *HST* (*F160W*), mid-infrared IRAC-Spitzer data, multi-object spectroscopy with ESO-FORS2, and Gemini-GMOS instruments. We have complemented those data with the spectroscopic catalogue from Momcheva et al. (2015) who spectroscopically measured the redshift of galaxies distant by up to 15 arcmin from WFI2033–4723 down to  $i \approx 21.5$ . In addition we have also used the exceptional capabilities of the ESO-MUSE integral-field spectrograph to derive spectroscopic redshift of the object closer in projection from the lens (with a projected distance as large as 30 arcsec from WFI2033–4723), but also to obtain velocity-dispersions of the brightest galaxies susceptible to produce high-order perturbation of the lens potential. With 64 galaxies having a confirmed redshift within a radius of 2 arcmin from the lens, we double the number of systems with a measured spectroscopic redshift in the direct vicinity of WFI2033–4723.

Our main results are the following:

(i) We have gathered a catalogue of 366 galaxies with confirmed spectroscopic redshifts in the FOV of WFI2033–4723. In addition, we have tentative redshift measurements for 24 galaxies, and 79 objects for which no redshift could be measured. We also spectroscopically identify 110 stars in the FOV.

(ii) We used the same methodology as Sluse et al. (2017) to identify groups of more than 5 (10) galaxies located within 6 (15) arcmin from the lensing galaxy. This selection does not aim at identifying all the groups along the line of sight, but those that are more susceptible affecting cosmological inference with the time-delay method, namely small groups close in projection from the lens, and/or more massive groups/clusters located farther away. Nine group candidates fulfilling those criteria were found, but two of them are likely to be spurious identifications. In particular, a0-3 (see Table 3) has fewer than two galaxies appearing in projection within one virial radius from its centroid. Another group candidate, a8, shows a bimodal redshift distribution unlikely to be associated with a single group, but our algorithm is unsuccessful in identifying this overdensity as two separated groups, or one group + isolated galaxies.

(iii) We confirm earlier findings that the main lensing galaxy is part of a large group at  $\bar{z}_{\text{group}} = 0.6588$  (Morgan et al. 2004; Wilson et al. 2016), for which we derive  $\sigma_{\text{los}} = 500 \pm 80 \text{ km s}^{-1}$ . The number of spectroscopically confirmed members has increased by 30 per cent owing to this work, and is now reaching 22 galaxies. The lensing galaxy is the seventh brightest galaxy of the group and is therefore suspected not to lie at the centre of its host halo.

(iv) Following McCully et al. (2017), we have calculated the flexion shift  $\Delta_{3x}$  to identify the galaxies/galaxy groups along the line-of-sight most susceptible to produce high-order perturbation of the lensing potential. Two groups may require to be included explicitly in the lens models: The group a3 at  $\bar{z}_{\text{group}} = 0.4956$ , for which we identified 13 group members, and the group a5 which hosts the lensing galaxy. In addition, three galaxies (*G2*, *G7*, *G3*, see Fig. 3) are susceptible to produce a non-negligible high-order perturbation of the main lens gravitational potential. Owing to our MUSE spectroscopic data, we have been able to measure the velocity-dispersion for these galaxies, which are used as a prior for the lens-modelling of WFI2033–4723 presented in Rusu et al. (2019).

(v) We measure the velocity-dispersion of the lens to be  $\sigma_{\text{los}} = 250^{+15}_{-21} \text{ km s}^{-1}$ .

These results are used by Rusu et al. (2019) to account for the main perturbers explicitly in the mass modelling of

WFI2033–4723, quantify the statistical contribution to the main lens potential of galaxies along the line of sight, and constrain  $H_0$  from the time delay measured in that lens system (Bonvin et al. 2019). Wong et al. (2019) present the constraints on various cosmological parameters combining the H0LiCOW lenses analysed so far.

## ACKNOWLEDGEMENTS

We thank Malte Tewes for his contribution to the H0LiCOW project. CER wishes to thank A. Tomczak for providing the PSF matching code. H0LiCOW and COSMOGRAIL are made possible thanks to the continuous work of all observers and technical staff obtaining the monitoring observations, in particular at the Swiss Euler telescope at La Silla Observatory. Euler is supported by the Swiss National Science Foundation. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 787886). This work was supported by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. TT thanks the Packard Foundation for generous support through a Packard Research Fellowship, the NSF for funding through NSF grant AST-1450141, ‘Collaborative Research: Accurate cosmology with strong gravitational lens time-delays’. SHS acknowledges support from the Max Planck Society through the Max Planck Research Group. KCW is supported in part by an EACOA Fellowship awarded by the East Asia Core Observatories Association, which consists of the Academia Sinica Institute of Astronomy and Astrophysics, the National Astronomical Observatory of Japan, the National Astronomical Observatories of the Chinese Academy of Sciences, and the Korea Astronomy and Space Science Institute. S.H. acknowledges support by the DFG cluster of excellence ‘Origin and Structure of the Universe’ ([www.universe-cluster.de](http://www.universe-cluster.de)). CER and CDF were funded through the NSF grant AST-1312329, ‘Collaborative Research: Accurate cosmology with strong gravitational lens time-delays’. AJS acknowledges support by NASA through STSCI grant HST-GO-15320. PJM acknowledges support from the U.S. Department of Energy under contract number DE-AC02-76SF00515. LVEK is supported in part through an NWO-VICI career grant (project number 639.043.308).

Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programme(s) 091.A-0642(A) (PI: Sluse), and 074.A-0302(A) (PI: Rix), 60.A-9306(A), 097.A-0454(A) (PI: Sluse), 090.A-0531(A) (PI: Fassnacht). Based on observations obtained at the Gemini Observatory (PID: GS-2013A-Q-2, PI: Treu), which is operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), and Ministério da Ciência, Tecnologia e Inovação (Brazil).

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell



Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft, and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2015–71825, ESP2015–66861, FPA2015–68048, SEV-2016–0588, SEV-2016–0597, and MDM-2015–0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007–2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) e-Universe (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

Based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. These observations are associated with program #12889. Support for program #12889 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. This work is based in part on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion

Laboratory, California Institute of Technology under a contract with NASA. This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

This work made extensive use of TOPCAT (Taylor 2014), and of Astropy, a community-developed core PYTHON package for Astronomy (The Astropy Collaboration 2013).

## REFERENCES

- Abbott T. M. C. et al., 2018, *ApJS*, 239, 18  
 Appenzeller I. et al., 1998, *The Messenger*, 94, 1  
 Arnouts S. et al., 2002, *MNRAS*, 329, 355  
 Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Bundy K., Moustakas L. A., 2009, *ApJ*, 705, 1099  
 Bacon R. et al., 2010, in *Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III*. SPIE, Bellingham, p. 773508  
 Beers T. C., Flynn K., Gebhardt K., 1990, *AJ*, 100, 32  
 Benítez N., 2000, *ApJ*, 536, 571  
 Bertin E., 2006, in *Gabriel C., Arviset C., Ponz D., Enrique S., eds, ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV*. Astron. Soc. Pac., San Francisco, p. 112  
 Bertin E., Arnouts S., 1996, *A&AS*, 117, 393  
 Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in *Bohlender D. A., Durand D., Handley T. H., eds, Astronomical Society of the Pacific Conference Series Vol. 281, Astronomical Data Analysis Software and Systems XI*. Astron. Soc. Pac., San Francisco, p. 228  
 Birrer S., Amara A., 2018, *Phys. Dark Universe*, 22, 189  
 Birrer S., Amara A., Refregier A., 2015, *ApJ*, 813, 102  
 Bonvin V., Tewes M., Courbin F., Kuntzer T., Sluse D., Meylan G., 2016, *A&A*, 585, A88  
 Bonvin V., Millon M., Chan J. H. -H., 2019, *A&A*, 629, 97  
 Braibant L., Hutsemékers D., Sluse D., Anguita T., García-Vergara C. J., 2014, *A&A*, 565, L11  
 Brammer G. B., van Dokkum P. G., Coppi P., 2008, *ApJ*, 686, 1503  
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000  
 Carrasco Kind M., Drlica-Wagner A., Koziol A. M. G., Petravick D., 2018, *JOSS*, 4, 1022  
 Chabrier G., 2003, *PASP*, 115, 763  
 Collett T. E. et al., 2013, *MNRAS*, 432, 679  
 Drake A. B. et al., 2017, *MNRAS*, 471, 267  
 Eigenbrod A., Courbin F., Meylan G., Vuissoz C., Magain P., 2006, *A&A*, 451, 759  
 Erben T. et al., 2013, *MNRAS*, 433, 2545  
 Fassnacht C. D., Koopmans L. V. E., Wong K. C., 2011, *MNRAS*, 410, 2167  
 Fazio G. G. et al., 2004, *ApJS*, 154, 10  
 Flaugher B. et al., 2015, *AJ*, 150, 150  
 Freudling W., Romaniello M., Bramich D. M., Ballester P., Forchi V., García-Dabó C. E., Moehler S., Neeser M. J., 2013, *A&A*, 559, A96  
 Gawiser E. et al., 2006, *ApJS*, 162, 1  
 Hildebrandt H. et al., 2010, *A&A*, 523, A31  
 Hildebrandt H. et al., 2012, *MNRAS*, 421, 2355  
 Hook I. M., Jørgensen I., Allington-Smith J. R., Davies R. L., Metcalfe N., Murowinski R. G., Crampton D., 2004, *PASP*, 116, 425  
 Hoshino H. et al., 2015, *MNRAS*, 452, 998  
 Ilbert O. et al., 2010, *ApJ*, 709, 644  
 Kissler-Patig M. et al., 2008, *A&A*, 491, 941  
 Koopmans L. V. E., 2004, in *Detmar J., Klein U., Salucci, P., eds, Proc. Sci., Conf.: Baryons in Dark Matter Haloes, 5-9 October 2004, SISSA, Novigrad, Croatia, preprint (astro-ph/0412596)*

Labbé I. et al., 2003, *AJ*, 125, 1107

Markwardt C. B., 2009, in Bohlender D. A., Durand D., Dowler P., eds, ASP Conf. Ser. Vol. 411, *Astronomical Data Analysis Software and Systems XV III*. Astron. Soc. Pac., San Francisco, p. 251

McCully C., Keeton C. R., Wong K. C., Zabludoff A. I., 2014, *MNRAS*, 443, 3631

McCully C., Keeton C. R., Wong K. C., Zabludoff A. I., 2017, *ApJ*, 836, 141

Merlin E. et al., 2015, *A&A*, 582, A15

Moffat A. F. J., 1969, *A&A*, 3, 455

Momcheva I. G., Williams K. A., Cool R. J., Keeton C. R., Zabludoff A. I., 2015, *ApJS*, 219, 29

Morgan N. D., Caldwell J. A. R., Schechter P. L., Dressler A., Egami E., Rix H.-W., 2004, *AJ*, 127, 2617

Morganson E. et al., 2018, *PASP*, 130, 074501

Muñoz R. P. et al., 2013, *A&A*, 552, A80

Piqueras L., Conseil S., Shepherd M., Bacon R., Leclercq F., Richard J., 2017, preprint ([arXiv:1710.03554](https://arxiv.org/abs/1710.03554))

Pirard J.-F. et al., 2004, in Moorwood A. F. M., Iye M., eds, Proc. SPIE Conf. Ser. Vol. 5492, *Ground-based Instrumentation for Astronomy*. SPIE, Bellingham, p. 1763

Quadri R. et al., 2007, *AJ*, 134, 1103

Richard J., R. B., J. V., 2017, MUSE Pipeline User Manual, VLT-MAN-ESO-261650, Issue 7.0. ESO, 7.0 edn, [https://www.eso.org/sci/activities/vltsv/muse/ESO-261650\\_7\\_MUSE\\_User\\_Manual.pdf](https://www.eso.org/sci/activities/vltsv/muse/ESO-261650_7_MUSE_User_Manual.pdf)

Robotham A. S. G. et al., 2011, *MNRAS*, 416, 2640

Rusu C. E. et al., 2017, *MNRAS*, 467, 4220

Rusu C. E. et al., 2019, *MNRAS*, preprint ([arXiv: 1907.04869](https://arxiv.org/abs/1907.04869))

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525

Shajib A. J., Treu T., Agnello A., 2018, *MNRAS*, 473, 210

Shen S., Yang X., Mo H., van den Bosch F., More S., 2014, *ApJ*, 782, 23

Skrutskie M. F. et al., 2006, *AJ*, 131, 1163

Sluse D., Hutsemékers D., Courbin F., Meylan G., Wambsganss J., 2012, *A&A*, 544, A62

Sluse D., Hutsemékers D., Anguita T., Braibant L., Riaud P., 2015, *A&A*, 582, A109

Sluse D. et al., 2017, *MNRAS*, 470, 4838

Sonnenfeld A., Treu T., Gavazzi R., Marshall P. J., Auger M. W., Suyu S. H., Koopmans L. V. E., Bolton A. S., 2012, *ApJ*, 752, 163

Soto K. T., Lilly S. J., Bacon R., Richard J., Conseil S., 2016, *MNRAS*, 458, 3210

Suyu S. H., Halkola A., 2010, *A&A*, 524, A94

Suyu S. H., Marshall P. J., Auger M. W., Hilbert S., Blandford R. D., Koopmans L. V. E., Fassnacht C. D., Treu T., 2010, *ApJ*, 711, 201

Suyu S. H. et al., 2012, *ApJ*, 750, 10

Suyu S. H. et al., 2017, *MNRAS*, 468, 2590

Taylor M. B., 2014, in Manset N., Forshay P., eds, *Astronomical Society of the Pacific Conference Series Vol. 485, Astronomical Data Analysis Software and Systems XXIII*. Astron. Soc. Pac., San Francisco, p. 257

The Astropy Collaboration, 2013, *A&A*, 558, A33

Treu T., Koopmans L. V. E., 2002a, *MNRAS*, 337, 6

Treu T., Koopmans L. V. E., 2002b, *ApJ*, 575, 87

Valdes F., Gupta R., Rose J. A., Singh H. P., Bell D. J., 2004, *ApJS*, 152, 251

Vuissoz C. et al., 2008, *A&A*, 488, 481

Weilbacher P. M., Streicher O., Urrutia T., Jarno A., Pécontal-Rousset A., Bacon R., Böhm P., 2012, in *Software and Cyberinfrastructure for Astronomy II*, Proceedings of SPIE, p. 84510B

Weilbacher P. M., Streicher O., Urrutia T., Pécontal-Rousset A., Jarno A., Bacon R., 2015, *ASPC*, 485, 451

Wilman D. J., Balogh M. L., Bower R. G., Mulchaey J. S., Oemler A., Carlberg R. G., Morris S. L., Whitaker R. J., 2005, *MNRAS*, 358, 71

Wilson M. L., Zabludoff A. I., Ammons S. M., Momcheva I. G., Williams K. A., Keeton C. R., 2016, *ApJ*, 833, 194

Wong K. C. et al., 2018, *ApJ*, 867, 107

Wong K. C. et al., 2019, *MNRAS*, preprint ([arXiv: 1907.04869](https://arxiv.org/abs/1907.04869))

Zahid J., Geller M., Fabricant D., Hwang H. S., 2016, *ApJ*, 832, 203

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

**Table 2.** Spectroscopic redshift catalogue.

**Table C2.** Photometric properties of the  $i \leq 23$  galaxies inside 120 arcsec of WFI2033–4723.

**Table C3.** Measured redshifts and stellar masses of the galaxies in Table C2.

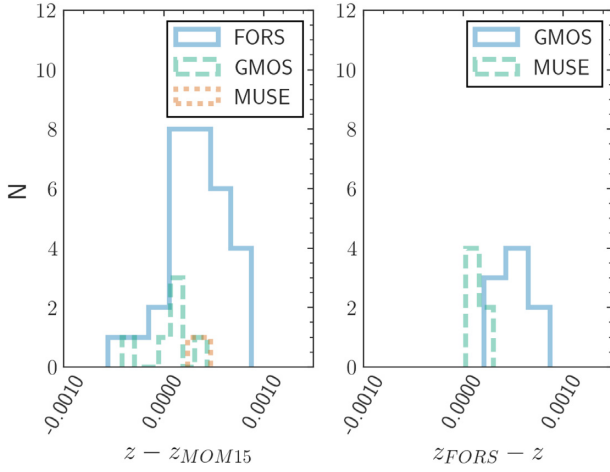
Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: COMPARISON WITH LITERATURE REDSHIFTS

Independent redshift measurements are available for a small sub-sample of objects. On one hand, we have compared the 38 robust galaxy measurements present in our catalogue and in MOM15 catalogue.<sup>9</sup> The results (split by instrument), are shown in the left-hand panel of Fig. A1. We find a median redshift offset with MOM15 of  $\delta z = 3.6 \times 10^{-4} \pm 8.5 \times 10^{-6}$  (standard error on the mean; stde), compatible with what we observed for HE 0435–1223. Since the width of the observed  $\delta z$  distribution is compatible with the median formal error on the redshift, we conclude that, while GMOS and MOM15 redshifts are mutually compatible, the offset between FORS and MOM15 is real. On the other hand, we have compared redshifts for targets in common in the FORS, GMOS, and MUSE catalogues (right-hand panel of Fig. A1). There are ine galaxies with robust redshift measurements in common between the GMOS and FORS catalogues. The distribution of redshift differences between the two catalogues is centred on  $\delta z = z_{\text{FORS}} - z_{\text{GMOS}} \sim 0.0005 \pm 0.00007$  (stde). There are six galaxies in common between the MUSE and the FORS samples, and we measure  $\delta z = z_{\text{MUSE}} - z_{\text{FORS}} \sim -0.00009 \pm 0.00005$  (stde). Despite the smaller sample of galaxies, those results support the results found by comparing our results to MOM15, namely that FORS and MUSE redshifts are mutually compatible but offset by about  $\delta z \sim 0.0004$  compared to MOM15 and GMOS.

These comparisons suggest that there are unaccounted systematic errors between the catalogues, but it is difficult to trace their origin. This has no impact on our analysis provided that we apply a systematic correction to match redshifts of all the catalogues. Similarly to the approach followed for HE 0435–1223, we have decided to correct the ESO-based data (i.e. MUSE and FORS). We apply for our analysis a systematic correction of  $\delta z = -3.6 \times 10^{-4}$  to the FORS and MUSE redshifts. Such an offset corresponds to  $\sim 1$  FORS pixel  $\sim 3.3 \text{ \AA}$  in our wavelength calibration (i.e. about five times the error on the wavelength calibration derived along the reduction). This translates into a velocity offset  $\delta v \sim 120 \text{ km s}^{-1}$ . In addition, we have quadratically added an error

<sup>9</sup>We only consider objects with the same redshift and with flags 3 and 4 in the MOM15, i.e. we exclude objects that are not new measurements from MOM15 but included in their catalogue.



**Figure A1.** Left-hand panel: distribution of the difference of redshifts between our spectra obtained with three different instruments (FORS, GMOS, MUSE) and MOM15 (i.e.  $\delta z = z_i - z_{\text{MOM15}}$ , where  $i = [\text{FORS}, \text{GMOS}, \text{MUSE}]$ ). Right-hand panel: Distribution of the difference of redshifts between FORS and GMOS / MUSE.

$\sigma_z^{\text{sys}}$  to the redshifts measured with different instruments. Based on the above comparisons, we have considered  $\sigma_z^{\text{sys}} = 0.0002$  for FORS, MUSE, and GMOS, and  $\sigma_z^{\text{sys}} = 0.0001$  for MOM15 data. This systematic error accounts for the uncertainty on the

offset between pairs of catalogues, and for a possible systematic error associated with the barycentric correction (our redshifts include a barycentric correction – smaller than the systematic offset  $\delta z$  – but we do not know if this is the case of MOM15 redshifts). Finally, we note that we do not account for the displacement of our Galaxy w.r.t. the Cosmic Microwave Background, which for WFI 2033–4723 implies an additional correction  $\delta z \sim 0.00053$  (that would be identical for all the catalogues). The impact of this correction on the cosmological analysis is yet negligible relative to other sources of uncertainties, and can otherwise be accounted for explicitly.

In addition, there are also four objects for which our spectra do not support the redshift measurement published by Momcheva et al. (2015). There is one object for which we propose to revise the redshift measurement. The discrepancy is small ( $\delta z = 0.04$ ) but significant. Since there is a nearby, but fainter galaxy 3.0 arcsec away from the target, we cannot fully exclude that the redshifts correspond to two different objects, but this seems unlikely. For three other targets, we could not measure any redshift, while Momcheva et al. (2015) publish a robust measurement. Since the SNR of our spectra is on average better, we suggest that there could be a potential mis-identification in Momcheva et al. (2015). Therefore, we choose for this analysis to set the quality flag to 2 when there is no confirmation of the published redshift with our new spectrum, and to 1 in case of possible support. Table A1 lists the coordinates and IDs of the galaxies with potentially discrepant redshifts.

**Table A1.** Objects with significantly different redshifts in Momcheva et al. (2015) and in our catalogue. The last column comments on the reason of likely mis-identification.

(RA, DEC) (deg)	ID-MOM	ID	$z_{\text{MOM}} (\sigma_z)$	$z (\sigma_z)$	Note
(308.417 400, −47.411 13)	13198	354	0.3540 (3.0e−4)	0.39485 (9.7e−5)	Unknown. Nearest galaxy is 3 arcsec
(308.390 047, −47.397 31)	11787	487	0.4997 (3.0e−4)	–	Unknown. Flagged changed to 2
(308.388 206, −47.391 50)	12266	1044	0.9701 (3.0e−4)	–	Unknown. Flagged changed to 2
(308.4031, −47.343 67)	12814	607	0.7514 (2.3e−4)	–	Unknown. Flagged changed to 1



## APPENDIX B: GROUPS FOUND WHEN USING A LUMINOSITY WEIGHTING SCHEME

As a comparison, we applied our group detection algorithm using a luminosity weighting scheme to calculate the group centroid. Table B1 displays the properties of the groups identified following this methodology. Groups at the same redshift and with similar velocity-dispersion as the fiducial ones are identified, except the group a0-1 (Table B1). On the other hand, the group candidates

identified with that weighting scheme seem to be less physically plausible, in particular because they often host fewer galaxies within one virial radius than in the fiducial case. This led to the flagging of the groups b5 ( $\bar{z}_{\text{group}} = 0.3059$ ) and a0-2 ( $\bar{z}_{\text{group}} = 0.3867$ ) as spurious group candidates. The group centroid is compatible with the one reported in Table 3 within  $3\sigma$ . The group centroid is generally found to be closer to a group galaxy in the fiducial case. In addition, the brightest galaxy is never found to be the closest to the centroid, even when luminosity weighting is used.

**Table B1.** Properties of the groups identified in the FOV of WFI 2033–4723. Same as Table 3 but considering a luminosity weighted centroid. The compact group a2-1 is not identified when we follow that methodology. The columns are the group redshift, the number of spectroscopically identified galaxies in the group, the group intrinsic velocity-dispersion (rounded to the nearest  $10 \text{ km s}^{-1}$ ) and  $1\sigma$  standard deviation from bootstrap, the group centroid, bootstrap error on the centroid, projected distance of the centroid to the lens, median flexion shift  $\log(\Delta_{3x}(\text{arcsec}))$  and  $1\sigma$  standard deviation from bootstrapping (Section 5). The last column indicates for which field a peak of more than five galaxies is detected in redshift space. The properties we display correspond to the FOV marked in bold.

ID	$\bar{z}_{\text{group}}$	$N$	$\sigma_{\text{int}} \pm \text{err}$ ( $\text{km s}^{-1}$ )	$\alpha_{\text{ctr}}, \delta_{\text{ctr}}$ (deg)	$\text{err}(\alpha_{\text{ctr}}, \delta_{\text{ctr}})$ (arcsec)	$\Delta\theta$ (arcsec)	$\log(\Delta_{3x}) \pm \text{err}$ [log(arcsec)]	FOV (arcmin)
b5	0.3059 <sup>a</sup>	10	490 (90)	308.505 394 18, −47.371 613 10	251.0, 145.1	212.7	−6.07 ± 0.63	900
a0-2	0.3867 <sup>a</sup>	8	100 (20)	308.445 701 76, −47.366 956 69	87.9, 20.4	113.8	−7.68 ± 0.51	360
a0-3	0.3999 <sup>a,b</sup>	12	380 (70)	308.342 706 00, −47.336 712 75	87.6, 77.7	292.5	−7.04 ± 0.44	360
a2	0.4436	6	150 (40)	308.472 528 46, −47.391 114 78	19.3, 25.7	115.5	−7.18 ± 0.46	360
a3	0.4956	13	520 (110)	308.428 142 92, −47.383 721 92	67.3, 66.6	42.8	−4.43 ± 0.95	360
a5	0.6588	22	500 (90)	308.445 125 58, −47.372 585 09	36.0, 20.2	95.2	−4.87 ± 0.43	<b>360</b> , 900
a8	0.6796 <sup>b</sup>	11	610 (200)	308.425 310 59, −47.393 185 38	71.5, 25.2	8.3	−3.71 ± 1.14	<b>360</b> , 900
a9	0.6889	4	190 (90)	308.413 903 09, −47.413 583 72	22.9, 15.3	71.1	−6.37 ± 2.20	360

<sup>a</sup>Likely spurious. <sup>b</sup>No luminosity weighting due to some group members missing magnitude measurement. <sup>c</sup>Results from the breakdown of a larger multimodal group candidate.

## APPENDIX C: MASKING FOR VELOCITY-DISPERSION MEASUREMENT

Two different masking schemes have been used to measure the velocity-dispersion of  $G$ . A first mask consists of regions identified as being susceptible to residual sky artefacts, as well as main telluric lines. The second mask contains the same features as the first mask but wavelengths bluer than 6600 Å are also masked out, as those regions are the most susceptible to uncertainties due to sky extraction. For G2, a third mask, containing only the known telluric absorption and the red edge of the spectrum ( $\lambda > 9100$  Å) has also been used.

We list in Table C1 the spectral range excluded for the velocity-dispersion measurements. This list contains known telluric absorption, as well as spectral ranges for which the data suggested sky variability larger than expectation based on the variance frame. To identify those regions, we proceeded as follows: we extract a spectrum in an aperture of the same size as the aperture used for the galaxies, for 10 different empty regions of the FOV. We then compare the observed variation of the spectrum, as a function of wavelength, to the median variance estimated based on the variance spectrum. We have considered that regions (limited to maximum 50 consecutive pixels) that deviate by more than three times the median standard deviation are the most susceptible to be poorly sky subtracted, and therefore are masked out. Similar regions are found for SV and P97 data, suggesting that those regions need effectively to be masked in addition to the spectral bands containing telluric lines. In addition to those regions, we have also added the reddest edge of the spectrum whose absolute level varies by more than 100 per cent between individual data-cubes in the SV data.

**Table C1.** Sky bands masked out for the velocity-dispersion measurements.  $\lambda_1$  and  $\lambda_2$  correspond respectively to the beginning and end of the wavelength band. The third column indicates if the band is a region known to be affected by telluric absorption or identified based on the study of the sky variability. For clarity, we have not merged overlapping masked regions.

$\lambda_1$ (Å)	$\lambda_2$ (Å)	Note
4750	4780	Telluric
5573	5584	Variable
5885	5960	Telluric
6295	6307	Variable
6359	6369	Variable
7716	7782	Variable
7791	7857	Variable
7800	7920	Telluric
7910	7968	Variable
7990	8030	Variable
8341	8405	Variable
8426	8471	Variable
8758	8790	Telluric
8762	8829	Variable
8820	8850	Telluric
8826	8892	Variable
8885	8895	Telluric
8902	8965	Variable
8910	8922	Telluric
8997	9056	Variable
9100	9344	Red edge

**Table C2.** Photometric properties of the  $i \leq 23$  galaxies inside 120 arcsec of WFI 2033–4723. (The complete table is provided as supporting information.)

RA	DEC	$u$	$g$	$r$	$i$	$z$	$Y$	$J$	$H$	$K$	3.6 $\mu$ m	4.5 $\mu$ m	5.7 $\mu$ m	7.9 $\mu$ m
308.421 01	−47.363 79	23.42 ± 0.07	23.65 ± 0.13	23.01 ± 0.09	22.44 ± 0.09	21.85 ± 0.15	21.76 ± 0.21	21.34 ± 0.16	21.06 ± 0.17	20.83 ± 0.11	20.96 ± 0.01	21.11 ± 0.02	21.86 ± 0.18	20.68 ± 0.13
308.411 23	−47.363 96	22.30 ± 0.04	21.38 ± 0.02	20.10 ± 0.01	19.55 ± 0.01	19.30 ± 0.02	19.05 ± 0.04	18.83 ± 0.03	18.44 ± 0.03	18.18 ± 0.02	18.88 ± 0.00	18.96 ± 0.00	19.40 ± 0.02	18.73 ± 0.02
308.440 54	−47.365 04	25.05 ± 0.15	24.53 ± 0.21	22.99 ± 0.06	21.76 ± 0.09	21.76 ± 0.09	21.30 ± 0.14	20.85 ± 0.07	20.39 ± 0.07	20.17 ± 0.06	20.03 ± 0.01	20.49 ± 0.01	20.55 ± 0.04	21.57 ± 0.22
308.447 39	−47.366 05	24.09 ± 0.09	24.09 ± 0.16	23.39 ± 0.09	22.83 ± 0.09	22.47 ± 0.17	—	22.34 ± 0.20	—	—	22.17 ± 0.03	22.72 ± 0.08	—	—
308.437 90	−47.366 45	23.43 ± 0.07	23.00 ± 0.07	22.42 ± 0.05	22.18 ± 0.06	22.26 ± 0.13	—	22.38 ± 0.18	21.43 ± 0.21	—	22.24 ± 0.04	22.30 ± 0.06	—	—
308.425 81	−47.368 86	22.61 ± 0.05	22.36 ± 0.05	21.46 ± 0.03	20.84 ± 0.03	20.68 ± 0.08	20.39 ± 0.10	20.65 ± 0.11	20.12 ± 0.11	20.10 ± 0.11	20.03 ± 0.01	20.49 ± 0.01	20.47 ± 0.05	20.99 ± 0.15
308.424 81	−47.369 99	24.15 ± 0.11	23.00 ± 0.06	21.46 ± 0.02	21.03 ± 0.03	20.88 ± 0.06	20.73 ± 0.10	20.67 ± 0.08	20.33 ± 0.10	20.37 ± 0.07	20.72 ± 0.01	20.89 ± 0.02	21.78 ± 0.13	21.52 ± 0.22
308.409 78	−47.367 09	23.30 ± 0.05	22.68 ± 0.05	22.25 ± 0.04	22.11 ± 0.05	21.94 ± 0.13	22.09 ± 0.13	21.76 ± 0.14	—	21.89 ± 0.21	23.51 ± 0.18	—	—	22.66 ± 0.72
308.440 13	−47.368 91	23.88 ± 0.06	23.89 ± 0.11	23.53 ± 0.08	22.93 ± 0.14	22.93 ± 0.14	21.72 ± 0.16	22.07 ± 0.16	22.05 ± 0.21	—	21.56 ± 0.02	21.90 ± 0.03	22.22 ± 0.18	21.28 ± 0.15
308.422 81	−47.369 42	23.53 ± 0.06	23.44 ± 0.09	22.87 ± 0.07	22.30 ± 0.07	22.53 ± 0.15	—	—	—	—	22.43 ± 0.04	22.87 ± 0.09	23.44 ± 0.65	21.62 ± 0.22
308.430 34	−47.369 87	24.01 ± 0.09	23.92 ± 0.13	23.19 ± 0.07	22.62 ± 0.08	22.77 ± 0.16	21.83 ± 0.18	—	—	—	22.41 ± 0.03	22.91 ± 0.09	—	—

*Note:* We report MAG-ISO magnitudes measured with SEXTRACTOR, but corrected to total magnitudes using MAG-AUTO<sub>1</sub> − MAG-ISO<sub>1</sub>, with detections in the combined  $ir$  bands. Reported magnitudes are corrected for galactic extinction following Schlafly, Finkbeiner & Davis (1998). For the IRAC channels, when computing the physical quantities in Table C3, additional uncertainties were added from the EAZY template error function (not included in this table).

**Table C3.** Measured redshifts and stellar masses of the galaxies in Table C2. (The complete table is provided as supporting information.)

RA	DEC	sep	zspec/EAZY	log $M_*$	RA	DEC	sep	zspec/EAZY	log $M_*$
308.421 01	−47.363 79	113.9	0.891 <sup>+0.129</sup> <sub>−0.124</sub>	10.40 <sup>+0.08</sup> <sub>−0.10</sub>	308.411 23	−47.363 96	117.9	0.388 <sup>+0.000</sup> <sub>−0.000</sub>	10.66 <sup>+0.03</sup> <sub>−0.04</sub>
308.440 54	−47.365 04	115.0	0.736 <sup>+0.051</sup> <sub>−0.054</sub>	10.45 <sup>+0.04</sup> <sub>−0.04</sub>	308.447 39	−47.366 05	118.2	0.724 <sup>+0.161</sup> <sub>−0.154</sub>	9.50 <sup>+0.46</sup> <sub>−0.50</sub>
308.437 90	−47.366 45	108.2	0.433 <sup>+0.127</sup> <sub>−0.148</sub>	7.99 <sup>+0.21</sup> <sub>−0.21</sub>	308.425 81	−47.368 86	95.1	0.651 <sup>+0.065</sup> <sub>−0.064</sub>	10.18 <sup>+0.05</sup> <sub>−0.06</sub>
308.424 81	−47.369 99	91.1	0.387 <sup>+0.000</sup> <sub>−0.000</sub>	9.71 <sup>+0.06</sup> <sub>−0.07</sub>	308.409 78	−47.367 09	108.3	0.228 <sup>+0.068</sup> <sub>−0.067</sub>	8.64 <sup>+0.54</sup> <sub>−0.62</sub>
308.440 13	−47.368 91	101.6	0.781 <sup>+0.146</sup> <sub>−0.127</sub>	9.89 <sup>+0.43</sup> <sub>−0.47</sub>	308.422 81	−47.369 42	93.3	0.642 <sup>+0.088</sup> <sub>−0.087</sub>	9.19 <sup>+0.58</sup> <sub>−0.66</sub>
308.430 34	−47.369 87	92.3	0.611 <sup>+0.111</sup> <sub>−0.114</sub>	9.07 <sup>+0.65</sup> <sub>−0.52</sub>	308.438 07	−47.370 58	94.2	0.537 <sup>+0.126</sup> <sub>−0.139</sub>	9.42 <sup>+0.08</sup> <sub>−0.09</sub>
308.412 18	−47.370 69	94.1	0.445 <sup>+0.000</sup> <sub>−0.000</sub>	9.26 <sup>+0.28</sup> <sub>−0.39</sub>	308.448 30	−47.372 09	100.6	0.762 <sup>+0.031</sup> <sub>−0.035</sub>	11.11 <sup>+0.06</sup> <sub>−0.06</sub>
308.456 96	−47.371 98	114.0	0.504 <sup>+0.088</sup> <sub>−0.092</sub>	9.24 <sup>+0.65</sup> <sub>−0.69</sub>	308.395 68	−47.372 50	109.3	0.668 <sup>+0.137</sup> <sub>−0.141</sub>	9.37 <sup>+0.52</sup> <sub>−0.56</sub>
308.448 98	−47.373 34	97.8	0.810 <sup>+0.119</sup> <sub>−0.116</sub>	10.20 <sup>+0.09</sup> <sub>−0.08</sub>	308.419 95	−47.373 93	78.0	0.566 <sup>+0.000</sup> <sub>−0.000</sub>	10.35 <sup>+0.08</sup> <sub>−0.09</sub>
308.421 87	−47.374 83	74.1	0.540 <sup>+0.154</sup> <sub>−0.165</sub>	9.13 <sup>+0.52</sup> <sub>−0.52</sub>	308.440 46	−47.375 07	81.6	0.494 <sup>+0.144</sup> <sub>−0.158</sub>	9.53 <sup>+0.11</sup> <sub>−0.14</sub>

*Note:* Where spectroscopic measurements (marked with null error bars) are not available, photometric redshift values (estimated with EAZY) correspond to the peak of the probability distributions, and logarithmic mass values (incorporating the IRAC photometry) correspond to the medians of the probability distributions estimated with LEPHARE, unless only the MASS\_BEST value was successfully computed by LEPHARE. Error bars mark the enclosed 68 per cent confidence regions.

<sup>1</sup>STAR Institute, Quartier Agora – Allée du six Août, 19c B-4000 Liège, Belgium

<sup>2</sup>Subaru Telescope, National Astronomical Observatory of Japan, 650 N Aohoku Pl, Hilo, HI 96720, USA

<sup>3</sup>Department of Physics, University of California, Davis, CA 95616, USA

<sup>4</sup>Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

<sup>5</sup>Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands

<sup>6</sup>CRAL, Observatoire de Lyon, F-69230 Saint-Genis-Laval, France

<sup>7</sup>Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

<sup>8</sup>European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching, Germany

<sup>9</sup>National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

<sup>10</sup>Max Planck Institute for Astrophysics, Karl-Schwarzschild-Strasse 1, D-85741 Garching, Germany

<sup>11</sup>Institute of Astronomy and Astrophysics, Academia Sinica, PO Box 23-141, Taipei 10617, Taiwan

<sup>12</sup>Physik-Department, Technische Universität München, James-Frank-Strasse 1, D-85748 Garching, Germany

<sup>13</sup>Department of Physics and Astronomy, PAB, 430 Portola Plaza, Box 951547, Los Angeles, CA 90095, USA

<sup>14</sup>Laboratoire d'Astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland

<sup>15</sup>Institute of Cosmology and Gravitation, University of Portsmouth, Burnaby Rd, Portsmouth PO1 3FX, UK

<sup>16</sup>Exzellenzcluster Universe, Boltzmannstr 2, D-85748 Garching, Germany

<sup>17</sup>Kapteyn Astronomical Institute, University of Groningen, PO Box 800, NL-9700 AV Groningen, the Netherlands

<sup>18</sup>Kavli Institute for Particle Astrophysics & Cosmology, PO Box 2450, Stanford University, Stanford, CA 94305, USA

<sup>19</sup>Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA

<sup>20</sup>Instituto de Física Teórica UAM/CSIC, Universidad Autónoma de Madrid, E-28049 Madrid, Spain

<sup>21</sup>CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France

<sup>22</sup>Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France

<sup>23</sup>Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK

<sup>24</sup>SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA

<sup>25</sup>Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), 28040 Madrid, Spain

<sup>26</sup>Laboratório Interinstitucional de e-Astronomia – LIneA, Rua Gal. José Cristino 77, Rio de Janeiro, RJ 20921-400, Brazil

<sup>27</sup>Department of Astronomy, University of Illinois at Urbana-Champaign, 1002 W. Green Street, Urbana, IL 61801, USA

<sup>28</sup>National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA

<sup>29</sup>Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain

<sup>30</sup>Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain

<sup>31</sup>Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, E-08193 Barcelona, Spain

<sup>32</sup>Observatório Nacional, Rua Gal. José Cristino 77, Rio de Janeiro, RJ 20921-400, Brazil

<sup>33</sup>Department of Physics, IIT Hyderabad, Kandi, Telangana 502285, India

<sup>34</sup>Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA

<sup>35</sup>Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

<sup>36</sup>Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA

<sup>37</sup>Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

<sup>38</sup>California Institute of Technology, 1200 East California Blvd, MC 249-17, Pasadena, CA 91125, USA

<sup>39</sup>Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 16, CH-8093 Zurich, Switzerland

<sup>40</sup>Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA

<sup>41</sup>Department of Physics, The Ohio State University, Columbus, OH 43210, USA

<sup>42</sup>Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

<sup>43</sup>Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

<sup>44</sup>Department of Astronomy/Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA

<sup>45</sup>Australian Astronomical Optics, Macquarie University, North Ryde, NSW 2113, Australia

<sup>46</sup>Departamento de Física Matemática, Instituto de Física, Universidade de São Paulo, CP 66318, São Paulo, SP 05314-970, Brazil

<sup>47</sup>George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA

<sup>48</sup>Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA



<sup>49</sup>*Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*

<sup>50</sup>*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra (Barcelona), Spain*

<sup>51</sup>*School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK*

<sup>52</sup>*Brandeis University, Physics Department, 415 South Street, Waltham, MA 02453, USA*

<sup>53</sup>*Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-859 Campinas, SP, Brazil*

<sup>54</sup>*Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA*

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.